TECHNISCHE UNIVERSITÄT
CHEMNITZ

COMPARING APPROXIMATIONS FOR RISK MEASURES RELATED TO SUMS OF CORRELATED LOGNORMAL RANDOM VARIABLES

MASTERARBEIT

Technische Universität Chemnitz

Fakultät für Mathematik

Integrierter Internationaler Master- und Promotionsstudiengang Mathematik

eingereicht von Maryna Karniychuk

geboren am 11. Juli 1978 in Cherepovets (Russland)

Betreuer: Dr.rer.nat. Matthias Richter (TU Chemnitz)

Tag der Verteidigung: 15. Dezember 2006

Contents

1 Introduction

1.1 Posing the problem

The control of the risks of banking operations is understood as an important public task. International organizations and professional associations are busy to develop new concepts to measure and control risks. The aim is to achieve a global standard in risk controlling. So-called downside risks have recently attracted a great interest in modern investment management. Downside risk (shortfall risk) is the probability that a special return level (target return, benchmark return) will not be exceeded. It is more consistent with the investor's perception of risk than the classical measures of risk, as variance or standard deviation. The use of variance or standard deviation as measures of risk is often criticized by investors because negative and positive returns are equally used to calculate these risk measures, while in the concept of asymmetrical risk measures, among which shortfall risk is the most elementary example, only undesired returns are used to calculate risk. Consequently, asymmetric risk measures are important alternatives to the variance. Well-known examples of downside risk measures are the following:

"Value-at-Risk or p-quantile risk measure"

$$Risk_1[X] = VaR_p(X-b) := -Q_p(X-b), \tag{1}$$

"Conditional Left Tail Expectation or mean excess loss"

$$Risk_2[X] = CLTE_p(X-b) := -E(X-b|X \le Q_p(X)), \tag{2}$$

where b is a fixed benchmark, p is a given probability $p \in (0,1)$ and $Q_p(X)$ is the p-quantile of X. Value-at-Risk is a general method to measure risk. VaR measures the worst loss under normal market conditions over a specific time interval at a given confidence level p. It answers the question: how much can I lose with $p\%$ probability over a pre-set horizon? Another way of expressing is that VaR is the lowest quantile of the potential losses that can occur during a specified time period. Conditional Left Tail Expectation estimates the expected value of the $b-X$ payments in the worst $p \cdot 100\%$ cases. In its turn Value-at-Risk yields the smallest value of this payments, it is indifferent towards the values which are above the level given by the quantile.

The main object of this thesis is the estimation of alternative risk measures presented above. The investigation of downside risk measures will be carried out on the basis of the Black-Scholes model of idealized financial market. The stock price $S(t)$ will be assumed to follow a geometric Brownian motion. Thus, consider an asset, which price process satisfies the stochastic differential equation

$$dS(t) = S(t)[\mu dt + \sigma dW_t], \; S(0) = S^0.$$

In the last formula μ designates the drift of the asset, while $\sigma > 0$ is the volatility of the asset and W_t is a standard Wiener process.

At time interval $[0,T]$ time points t_k $(k = 0, 1, ..., n)$ will be fixed, such that

$$0 = t_0 < t_1 < ... < t_{n-1} < t_n = T.$$

Consider the *terminal wealth problem*, when the decision maker will invest a given series of *positive* saving amounts α_k $(k = 0, ..., n-1)$ at the predetermined time points

t_k. Denote by $\varphi(t)$ and $V(t)$ the number of the security which is held at time point t and the wealth value at time instant t respectively. Then it holds:

$$\varphi(0) = \frac{\alpha_0}{S^0},$$

$$V(0) = \alpha_0 = S^0\varphi(0).$$

Consequently, the final wealth is defined as:

$$\begin{aligned} V &= V(T) = \varphi(T)S(T) \\ &= \left(\varphi(0) + \sum_{i=1}^{n}(\varphi(t_i) - \varphi(t_{i-1}))\right) S(T) \\ &= \left(\frac{\alpha_0}{S^0} + \sum_{i=1}^{n-1} \frac{\alpha_i}{S(t_i)}\right) S(T) = \sum_{i=0}^{n-1} \alpha_i \frac{S(T)}{S(t_i)}. \end{aligned}$$

Obviously, for $0 \leq t_i \leq T$

$$\frac{S(T)}{S(t_i)} = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)(T - t_i) + \sigma(W_T - W_{t_i})\right),$$

i.e. the ratio of final and intermediate values of the stock is a lognormally distributed random variable, it holds

$$\ln \frac{S(T)}{S(t_i)} \sim N\left(\left(\mu - \frac{\sigma^2}{2}\right)(T - t_i), \sigma^2(T - t_i)\right).$$

Henceforth from the reason of convenience the following notation will be introduced

$$Z(i) := \left(\mu - \frac{\sigma^2}{2}\right)(T - t_i) + \sigma(W_T - W_{t_i}).$$

$Z(i)$ is the random accumulation factor over the period $[t_i, t_n]$. It can be presented as a linear combination of the components of the random vector $(Z_0, Z_1, ..., Z_{n-1})$, where Z_i denotes a stochastic return over the period $[t_i, t_{i+1}]$, $i = 0, ..., n-1$. This implies

$$Z(i) = \sum_{j=i}^{n-1}\left(\left(\mu - \frac{\sigma^2}{2}\right)(t_{j+1} - t_j) + \sigma(W_{t_{j+1}} - W_{t_j})\right) = \sum_{j=i}^{n-1} Z_j.$$

Define for $i = 1, ..., n$

$$\widehat{Z}(i) := Z(n-i) = \sum_{j=n-i}^{n-1} Z_j = \sum_{k=1}^{i} Z_{n-k} = \sum_{k=1}^{i} \widehat{Z}_k,$$

where $\widehat{Z}_k$ are defined as

$$\widehat{Z}_k := Z_{n-k}, \qquad k = 1, ..., n.$$

$\widehat{Z}_k$ are independent normally distributed random variables with mean $(t_{n-k+1} - t_{n-k})(\mu - \frac{\sigma^2}{2})$ and variance $(t_{n-k+1} - t_{n-k})\sigma^2$.

Thus the final value is:

$$V = V(T) = \sum_{i=0}^{n-1} \alpha_i e^{Z(i)} = \sum_{k=1}^{n} \alpha_{n-k} e^{Z(n-k)}.$$

Consequently, the chain of transformations given above, i.e. simple rearrangement of the accumulation factors and payments, leads to the general representation of the final value V, which will be used throughout this paper:

$$V = \sum_{k=1}^{n} \widehat{\alpha}_k e^{\widehat{Z}(k)} \tag{3}$$

where $\widehat{Z}(k)$ are "permutated" accumulation factors, i.e.

$$\widehat{Z}(k) = Z(n-k), \qquad k = 1, ..., n$$

and $\widehat{\alpha}_k$ are "permutated" payments, i.e.

$$\widehat{\alpha}_k = \alpha_{n-k}, \qquad k = 1, ..., n.$$

It should be noted that the present value of a series of future deterministic payments at times t_k; $k = 1, ..., n$ can be also written in the form (3), where $\widehat{Z}(k)$ now denotes the random discount factor over the period $[t_0, t_k]$, and $\widehat{\alpha}_k$ is a saving amount at time point t_k. Thus, all the results presented in this thesis can be easily transferred to the case of stochastic present value.

Even though the assumption of mutual independence between the components of corresponding sums is a very convenient one, it is sometimes not realistic, as is seen in our situation. The random variable V defined in (3) will be a sum of non-independent lognormal random variables. Sums of lognormals frequently appear in a variety of situations including engineering and financial mathematics. As was already mentioned above, typical examples are present values of future cash-flows with stochastic (Gaussian) interest rates (see Dhaene et al. (2002b) or Vanduffel et al. (2005a)). The pricing of Asian options (see e.g. Vanmale et al. (2006), Simon et el. (2000) or Reynaerts et al.(2006)) and basket options (see Deelstra et al. (2004)) is related to the distributions of such sums.

To calculate the risk measures we are interested in the distribution function of a sum of random variables in the form (3). Unfortunately, there is no general explicit formula for the distribution of sums of lognormal random variables. That is why usually one has to use time consuming Monte Carlo simulations. Despite the increase of computational power, which is observed last years, the computational time remains a serious drawback of Monte Carlo simulations especially when one has to estimate very high values of quantiles (e.g. solvency capital of an insurance company can be determined as 99.95% quantile, which is extremely difficult to estimate within reasonable time of simulation). Therefore alternative solutions were proposed. Among them are moment matching methods for approximating the distribution function of V: lognormal approximation, which is widely used in practice, and reciprocal (inverse) Gamma approximation. Both techniques approximate the unknown distribution function by a given one such that the first two moments coincide.

Recently Dhaene et al. derived comonotonic upper bound, lower bound and "maximal variance" lower bound approximations for the distribution function of V. The aim of this thesis is to compare comonotonic approximations for computing risk measures related to a sum of correlated lognormal random variables with two well-known moment-matching approximations.

1.2 Basic concepts and notations

The fundamental concepts and notations presented below hold true throughout the thesis. The triple $\{\Omega, \mathcal{A}, P\}$ is a probability space. All random variables are defined on this probability space. For every random variable X a function $F_X(x)$ of a real argument x, known as a cumulative distribution function of the random variable, is defined in the context of thesis by $F_X(x) = P(X \leq x)$. The distribution function $F_X(x)$ of a random variable X is non-negative, non-decreasing and right-continuous function with the property that

$$F_X(-\infty) = \lim_{x \to -\infty} F_X(x) = 0;$$
$$F_X(+\infty) = \lim_{x \to +\infty} F_X(x) = 1.$$

In the sequel, the notation $\overline{F_X}(x)$ will be used for decumulative distribution function,

$$\overline{F_X}(x) = 1 - F_X(x).$$

The inverse of a distribution function is usually defined as a non-decreasing and left-continuous function, $F_X^{-1}(p)$ such that

$$F_X^{-1}(p) := \inf\{x \in \mathbb{R} | F_X(x) \geq p\}, \quad p \in [0, 1]$$

with $\inf \emptyset = +\infty$. For all $x \in \mathbb{R}$ and $p \in [0, 1]$ holds

$$F_X^{-1}(p) \leq x \Longleftrightarrow p \leq F_X(x). \tag{4}$$

This assertion can be proved as follows. Let assume that

$$F_X^{-1}(p) \leq x.$$

In view of the fact that the distribution function $F_X(x)$ of a random variable X is defined as a non-decreasing function, it follows that

$$F_X(F_X^{-1}(p)) \leq F_X(x).$$

From the definition of $F_X^{-1}(p)$ it follows immediately

$$F_X(F_X^{-1}(p)) \geq p.$$

Consequently,

$$p \leq F_X(F_X^{-1}(p)) \leq F_X(x).$$

Now assume that

$$p \leq F_X(x).$$

Then because of the definition of $F_X^{-1}(p)$ it holds clearly

$$F_X^{-1}(p) \leq x.$$

Clearly, that if $F_X(x) = p$ is true for some interval of values x, then any element of this interval can serve as the inverse of the distribution function. Thus, a more sophisticated definition will be introduced and used later on.

Definition 1.1: For any real $p \in [0,1]$, a possible choice for the inverse of $F_X(x)$ in p is any point in the closed interval

$$[\inf\{x \in \mathbb{R}|F_X(x) \geq p\},\ \sup\{x \in \mathbb{R}|F_X(x) \leq p\}],$$

where $\inf \emptyset = +\infty$, and $\sup \emptyset = -\infty$. Taking the left hand border of this interval to be the value of the inverse distribution function at p, we get $F_X^{-1}(p)$. Similarly we define

$$F_X^{-1+}(p) = \sup\{x \in \mathbb{R}|F_X(x) \leq p\}, \quad p \in [0,1]$$

as the right hand border of the interval. It is a non-decreasing and right-continuous function. Note that $F_X^{-1}(0) = -\infty$, $F_X^{-1+}(1) = +\infty$ and $F_X^{-1}(p)$, $F_X^{-1+}(p)$ are finite for all $p \in (0,1)$.

Definition 1.2: For $\alpha \in [0,1]$, we define the *α-mixed inverse* of F_X as follows:

$$F_X^{-1(\alpha)}(p) = \alpha F_X^{-1}(p) + (1-\alpha)F_X^{-1+}(p), \quad p \in (0,1),$$

which is a non-decreasing function.

Remark 1.1: This more sophisticated definition of the inverse of the cumulative distribution function makes it possible to choose a particular inverse distribution function with the property that the relation $F_X^{-1}(F_X(d)) = d$ holds for certain d.

The following two important relations between the inverse distribution functions of the random variables X and $g(X)$ for a monotone function g will be frequently used.

Remark 1.2: Let X and $g(X)$ be real-valued random variables, and let $0 < p < 1$.

(1) If g is non-decreasing and left-continuous, then

$$F_{g(X)}^{-1}(p) = g(F_X^{-1}(p)).$$

(2) If g is non-decreasing and right-continuous, then

$$F_{g(X)}^{-1+}(p) = g(F_X^{-1+}(p)).$$

The proof of Remark 1.2 (assertion 1) can be found in Dhaene et al. (2002a)

If should be emphasized that only values of p corresponding to horizontal segments of F_X lead to different values of $F_X^{-1}(p)$, $F_X^{-1+}(p)$ and $F_X^{-1(\alpha)}(p)$.

In the sequel, the following indicator function of the random event A is in use

$$I_A(\omega) = I_A = \begin{cases} 1; & \text{if } \omega \in A; \\ 0; & \text{if } \omega \notin A. \end{cases}$$

By $\Phi(x)$ the cumulative distribution function of standard normal distribution will be denoted. It is given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt.$$

In the context of this thesis a risk will be understood as a random variable. Sometimes, additionally the existence of some moments will be required. A risk measure summarizes the information contained in the distribution function of a random variable in one single number that quantifies the risk exposure in a way that is meaningful for the problem at hand.

Definition 1.3: The *risk measure* ρ is defined as the mapping ρ: $L \to \mathbb{R}$, where L is a subset of the set of random variables.

In other words the risk measure is defined as a mapping from the set of random variables representing the risks at hand to the real line. Random variables are usually considered as losses or payments, that have to be made. A negative outcome for the loss variable means that gain has occurred.

Another actuarial notion which will be frequently used throughout this thesis is the notion of stop-loss premium. Reinsurance treaties usually cover only part of the risk. *Stop-loss (re)insurance* covers the top part. It is defined as follows: if the loss is X (we assume that $X \geq 0$) the payments equals

$$(X-d)_+ := \max\{X-d, 0\} = \begin{cases} X-d; & \text{if } X > d; \\ 0; & \text{if } X \leq d. \end{cases}$$

We decided to use the "actuarily concept" where large values are "bad" and transform the results to our situation. The insurer retains a risk d (his retention) and lets the reinsurer pay for the remainder. In the reinsurance practice, the retention equals the maximum amount to be paid out for every single claim and d is called the priority. Why this coverage is called "stop-loss" is obvious: from the insurer's point of view, the loss stops at d.

In the context of thesis, the net premium $E[(X-d)_+]$ for a stop-loss contract will be understood under the notion of stop-loss premium.

Definition 1.4: A random vector $Y = (Y_1, ..., Y_n)^T$ is said to have a *multivariate normal distribution* with parameters μ, Σ if and only if for every vector b, the linear combination $b^T Y$ of the marginals Y_k has a univariate normal distribution with mean $b^T \mu$ and variance $b^T \Sigma b$.

2 Ordering random variables

The comonotonic bounds for sums of non-independent random variables, which are recently derived in the actuarial literature, are bounds in the terms of "convex order". Ordering random variables has been always the object of special interest for actuarial mathematicians. Several ordering concepts, such as stochastic dominance, stop-loss and convex order are introduced in literature.

Definition 2.1: Consider two random variables X and Y. Then X is said to precede Y in the *stochastic dominance sense*, notation $X \leq_{st} Y$, if and only if the distribution function of X always exceeds that of Y:

$$F_X(x) \geq F_Y(x), \quad -\infty < x < +\infty.$$

Another natural ordering concept in actuarial science is the stop-loss order.

Definition 2.2: Assume that the expectations of the random variables X and Y exist. Then X is said to precede Y in the *stop-loss order sense*, notation $X \leq_{sl} Y$, if and only if X has lower stop-loss premiums than Y:

$$E[(X-d)_+] \leq E[(Y-d)_+], \quad -\infty < d < +\infty$$

where $E[(X-d)_+]$, $E[(Y-d)_+]$ are stop-loss premiums with retention d of X and Y respectively.

Remark 2.1: Stop-loss order between random variables implies a corresponding ordering of their means; i.e.

$$X \leq_{sl} Y \Longrightarrow E[X] \leq E[Y].$$

The proof will be given later (see proof of Lemma 2.2).

Stop-loss order has many useful invariance properties. For example, it survives the operations of convolution and compounding on non-negative random variables (risks); stop-loss larger claims lead to increased ruin probability and higher zero-utility premiums for risk averse decision makers (for details see Kaas et al. (2001)). Risk X is preferred to Y either because it represents a smaller loss, or because it is less spread. Thus, random variable Y can be characterized as "less attractive" random variable.

It is a widely used trick in actuarial practice *to replace a random variable for which it is difficult to obtain the distribution function by a "less attractive" / "more dangerous" random variable with simpler structure*, i.e. by a random variable for which it is easier to obtain the distribution function. Now, taking into account Remark 2.1, confirming that stop-loss order between random variables preserves corresponding ordering of their means, it becomes intuitively clear, that the best approximation arises in the borderline case, i.e. when $E[X] = E[Y]$. This chain of consistent reasoning has led actuarial mathematicians to the concept of convex order.

In order to illustrate that convex order nicely suits the notion of dangerousness, the following definition of this concept will be given.

Definition 2.3: The random variable Y is said to be an upper bound for X in *convexity order*, notation $X \leq_{cx} Y$, if

$$E[X] = E[Y]$$

and

$$E[(X-d)_+] \leq E[(Y-d)_+]$$

for each value of d. Last definition implies that the financial loss of realizations exceeding a retention d, or stop-loss premium, is always larger for Y than for X and thus the variable Y is "more dangerous".

In the following lemma, two equivalent characterizations for convex order precedence are given.

Lemma 2.1: If the random variable X precedes Y in the convex order sense, i.e. $X \leq_{cx} Y$, then the following equivalent conditions hold:

a) $E[u(X)] \leq E[u(Y)]$ for each convex function u: $\mathbb{R} \longrightarrow \mathbb{R}$;
since convex functions take on their largest values in the tails, the variable Y is more likely to take on extremal values than the variable X and thus Y is "more dangerous".

b) $E[u(-X)] \geq E[u(-Y)]$ for each concave function u: $\mathbb{R} \longrightarrow \mathbb{R}$;
each risk averse decision maker prefers a loss X to a loss Y and thus the random variable Y is "more dangerous".

The proof that convex order implies ordered expectations of convex functions generally relies on the classical argument that any convex function can be obtained as the uniform limit of a sequence of piecewise linear functions. A simple proof will be given at the end of this section, as it is based on some of the results presented below.

From the view point of insurer, replicating random variable X with unknown distribution function by a variable Y with known distribution but larger in convex order sense is a cautious strategy. By the way, the characterization of stop-loss order in terms of utility functions is equivalent to $E[u(X)] \leq E[u(Y)]$ holding for all *non-decreasing* convex functions $u(\cdot)$, for which the expectations exist. Hence, it represents the common preferences of all risk averse decision makers. On the other hand, from Lemma 2.1(a) follows that convex order is the same as ordered expectations for *all* convex functions. This explains where the name convex order comes from. In the context of utility theory convex order represents the common preferences of all risk averse decision makers between random variables with equal mean.

Remark 2.2: It should be noted that taking $u(\cdot)$ as $u(x) = x^2$ in Lemma 2.1 it immediately follows that $Var[X] \leq Var[Y]$, whenever $X \leq_{cx} Y$.

Recalling the Definition 2.3 of the convex order precedence, an equivalent definition can be derived from the following relation

$$E[(X-d)_+] - E[(d-X)_+] = E[X] - d. \tag{5}$$

In particular,

$$X \leq_{cx} Y \Longleftrightarrow \begin{cases} E[X] = E[Y] \\ E[(d-X)_+] \leq E[(d-Y)_+] \qquad -\infty < d < \infty. \end{cases}$$

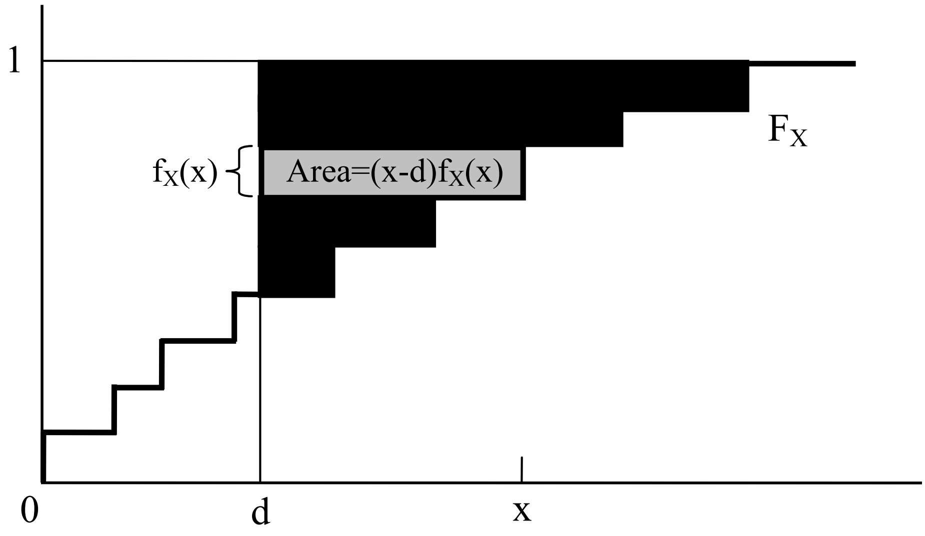

Fig. 1: Graphical derivation of stop-loss premium $E[X-d]_+$ for a discrete cumulative distribution function

The expected value of random variable X, if it exists, can be written as

$$E[X] = \int_{-\infty}^{0} x dF_X(x) - \int_{0}^{\infty} x d(1 - F_X(x)). \tag{6}$$

Taking into account, that random variables with finite means are considered, i.e.

$$\lim_{x\to\infty} x\overline{F_X}(x) = \lim_{x\to-\infty} xF_X(x) = 0,$$

partial integration of both terms in the expression for expected value given by (6) leads to

$$E[X] = -\int_{-\infty}^{0} F_X(x)dx + \int_{0}^{\infty} \overline{F_X}(x)dx. \tag{7}$$

Basing on the general Formula (7) for the expected value of a random variable X, it can be shown that in the continuous case, where $F_X(x)$ has $f_X(x)$ as its derivative, as well as in the discrete case, where $F_X(x)$ is a step function with a step $f_X(x)$ in x, the stop-loss premium is given by

$$E[(X-d)_+] = \begin{cases} \displaystyle\int_d^{\infty} (x-d) f_X(x)dx & \text{(continuous)} \\ \displaystyle\sum_{x>d} (x-d) f_X(x) & \text{(discrete)} \end{cases} = \int_d^{\infty} [1 - F_X(x)]dx. \tag{8}$$

A graphical "proof" for the discrete case is given by Fig. 1. The right hand side of the equation, i.e. the total shaded area enclosed by the graph of $F_X(x)$, the horizontal line at 1 and the vertical line at d, is divided into bars with a height $f_X(x)$ and a width $x-d$. It can be seen that the total area actually equals the left hand side of (8). The general case can be proved by partial integration

$$\begin{aligned} E[(X-d)_+] &= \int_d^\infty (x-d)dF_X(x) \\ &= -(x-d)[1-F_X(x)]\Big|_d^\infty + \int_d^\infty [1-F_X(x)]dx \\ &= \int_d^\infty \overline{F_X}(x)dx. \end{aligned} \tag{9}$$

That the integrated term vanishes for $x \to \infty$ is proved as follows: since $E[X] < \infty$, the integral $\int_0^\infty x dF_X(x)$ is convergent, and hence the "tails" tend to zero, so

$$x[1-F_X(x)] = x\int_x^\infty dF_X(t) \le \int_x^\infty t dF_X(t) \longrightarrow 0 \quad \text{for } x \longrightarrow \infty.$$

Consequently, the stop-loss premium with retention d can be considered as the weight of an upper tail of the distribution function of X, i.e. it is the surface between the cumulative distribution function F_X and constant function 1, from d on. Thus, stop-loss order entails uniformly heavier upper tails.

It can be shown by analogy that

$$\begin{aligned} E[(d-X)_+] &= \int_{-\infty}^{+\infty} (d-x)_+ dF_X(x) \\ &= \int_{-\infty}^{d} (d-x)dF_X(x) \\ &= (d-x)F_X(x)\Big|_{-\infty}^{d} + \int_{-\infty}^{d} F_X(x)dx \\ &= \int_{-\infty}^{d} F_X(x)dx. \end{aligned}$$

This means that $E[(d-X)_+]$ can be interpreted as the weight of a lower tail of X, i.e. it is the surface between the x-axis and the cumulative distribution function of X, from $-\infty$ to d. This implies that convex order can be characterized by uniformly heavier lower tails.

The quantity $E[(X-d)_+]$ represents the expected loss over d. Not only excessively large positive values of random variable are unattractive, but also negative ones (e.g. if positive values are assumed to be losses, hence negative values are actually gains, which might be also undesirable for the decision maker, for instance for tax reasons). Hence $E[(-X-t)_+]$ should be small too. So in this case, random variable X is preferred over Y if for all real $d=-t$ both

$$\begin{cases} E[(X-d)_+] \le E[(Y-d)_+] \quad \text{and} \\ E[(d-X)_+] \le E[(d-Y)_+]. \end{cases} \tag{10}$$

It can be shown that these two conditions are equivalent to $X \le_{cx} Y$.

Lemma 2.2: Conditions (10) hold if and only if $X \le_{cx} Y$.

Proof. "$\Longrightarrow$" Adding d to the expression (8) of stop-loss premium, and letting $d \to -\infty$ leads to $E[X] \leq E[Y]$. To be more precise

$$\begin{aligned} d + E[(X-d)_+] &= d + \int_d^{\infty} \overline{F_X}(x)dx \\ &= -\int_d^0 dx + \int_d^0 (1-F_X(x))dx + \int_0^{\infty} (1-F_X(x))dx \\ &= -\int_d^0 F_X(x)dx + \int_0^{\infty} \overline{F_X}(x)dx. \end{aligned}$$

It immediately follows from the expression (7) for the expectation of X that

$$\lim_{d \to -\infty} (d + E[(X-d)_+]) = E[X].$$

Consequently, from the first inequality in (10) follows $E[X] \leq E[Y]$. It should be emphasized that the result proved above is nothing else as the proof of Remark 2.1.

Further, subtracting d in the second set of equalities and letting $d \longrightarrow +\infty$ produces $E[X] \geq E[Y]$, hence $E[X] = E[Y]$. The equality of means together with the first set of inequalities from (10) leads to $X \leq_{cx} Y$ (see Definition 2.3).
"$\Longleftarrow$" Assume that $X \leq_{cx} Y$. According to the relation (5)

$$E[(X-d)_+] - E[(d-X)_+] = E[X] - d$$

and

$$E[(Y-d)_+] - E[(d-Y)_+] = E[Y] - d.$$

As $X \leq_{cx} Y \Longrightarrow E[X] = E[Y]$ and $E[(X-d)_+] \leq E[(Y-d)_+]$. This proves that $E[(d-X)_+] \leq E[(d-Y)_+]$. ⊠

In case $X \leq_{cx} Y$ the upper tails as well as lower tails of Y surpass that of X. Thus, extreme values are more typical for Y than for X. This observation also implies that $X \leq_{cx} Y$ is equivalent to $-X \leq_{cx} -Y$. Thus it does not matter whether random variables are interpreted as incomes or gains.

Finally, at the end of this section the proof of Lemma 2.1 is presented.

Proof of Lemma 2.1 (follows the ideas of Kaas et al. (2001)).
a) Consider the function $g(x) = u(x) - u(a) - (x-a)u'(a)$, with a some point where $u(\cdot)$ is differentiable. Assume that $X \leq_{cx} Y$. Since $E[X] = E[Y]$, the inequality $E[u(X)] \leq E[u(Y)]$ is equivalent to $E[g(X)] \leq E[g(Y)]$. As $g(a) = g'(a) = 0$, the integrated terms below vanish, so applying four times the method of integrating by parts one obtains

$$\begin{aligned} E[g(X)] &= \int_{-\infty}^{a} g(x)dF_X(x) - \int_a^{+\infty} g(x)d\overline{F_X}(x) \\ &= -\int_{-\infty}^{a} g'(x)F_X(x)dx + \int_a^{+\infty} g'(x)\overline{F_X}(x)dx \\ &= \int_{-\infty}^{a} E[(x-X)_+]dg'(x) + \int_a^{+\infty} E[(X-x)_+]dg'(x). \end{aligned}$$

From (10) immediately follows that $E[g(X)] \leq E[g(Y)]$, because by the convexity of $u(\cdot)$ and, consequently, $g(\cdot)$ we have $dg'(x) \geq 0$ for all x.
b) The assertion is obvious. It immediately follows from (a). ⊠

3 Comonotonicity

3.1 Comonotonic sets and random vectors

Many financial and actuarial applications are faced with the difficulty or impossibility to derive an analytical expression for the distribution of underlying stochastic quantity. In many cases, this difficulty arise from the presence of dependent components in this quantity. In current situation, in contrast to the quantities $\widehat{Z}_k$, the stochastic variables $\widehat{Z}(k)$ in (3) and, consequently, the payments $X_k := \widehat{\alpha}_k e^{\widehat{Z}(k)}$ are dependent, since they are constructed as successive series of the same sequence of independent variables. Note that for $k < l$ it holds $X_l = \frac{\widehat{\alpha}_l}{\widehat{\alpha}_k} X_k e^{(\widehat{Z}_{k+1}+...+\widehat{Z}_l)}$. So, in general, X_k and X_l have a large correlation in many random financial models.

In order to overcome this problem, Kaas et al. (2000) present bounds in convexity order that make use of the concept of comonotonic risks. This implies, that it might be helpful to replace the original sum V by a new sum, for which the components have the same marginal distributions as the components in the original sum, but with the *most "dangerous"* dependence structure (see Section 2 for details about convex order as the notion of dangerousness). The most "dangerous" or, saying mathematically, convex-largest sum will be obtained in the case when the random vector $(X_1, X_2, ..., X_n)$ with $X_i = \widehat{\alpha}_i e^{\widehat{Z}(i)}$, $i = 1, ..., n$ has the *comonotonic* distribution. The concept of comonotonicity was introduced by Yaari (1987) and Röell (1987) and has since then been playing a very important role in the economic theories of decision under risk and uncertainty. It should be noted that other characterizations of comonotonicity can be found in Denneberg (1994).

As the first step, define the comonotonicity of a set of real n-dimensional vectors in $\mathbb{R}^n$. Let x be the notation for the n-dimensional vector $(x_1, x_2, ..., x_n)$. The notation $x \leq y$ implies that $x_i \leq y_i$ for all $i = 1, 2, ..., n$, i.e. x and y are ordered componentwise.

Definition 3.1: The set $A \subseteq \mathbb{R}^n$ is referred to as *comonotonic* if for any x and y in A, either $x \leq y$ or $y \leq x$ holds, i.e. all components of the larger vector are at least the corresponding components of the other.

Notice, that a comonotonic set is a "thin" set. Since the upper left and lower right corners of a rectangle may not be in it, because of $x_1 < x_2$ but $y_1 > y_2$ (see Fig. 2), it must be a curve that is monotonically non-decreasing / non-increasing in each component (e.g. it might be a diagonal or a suitable curve connecting points (x_1, y_2) and (x_2, y_1)). Any subset of this curve is also comonotonic.

Definition 3.2: Denote the *(i, j)-projection* of set A in $\mathbb{R}^n$ by $A_{i,j}$

$$A_{i,j} = \{(x_i, x_j) | x \in A\}.$$

Lemma 3.1: *$A \subseteq \mathbb{R}^n$ is comonotonic if and only if $A_{i,j}$ is comonotonic for all $i, j \in \{1, 2, ..., n\}$, $i \neq j$.*

Proof. Straightforward.

Next define the notion of support of the n-dimensional random vector $X = (X_1, X_2, ..., X_n)$.

Definition 3.3: The smallest subset $A \subseteq \mathbb{R}^n$ with the property that $P[X \in A] = 1$ is called a *support* of X.

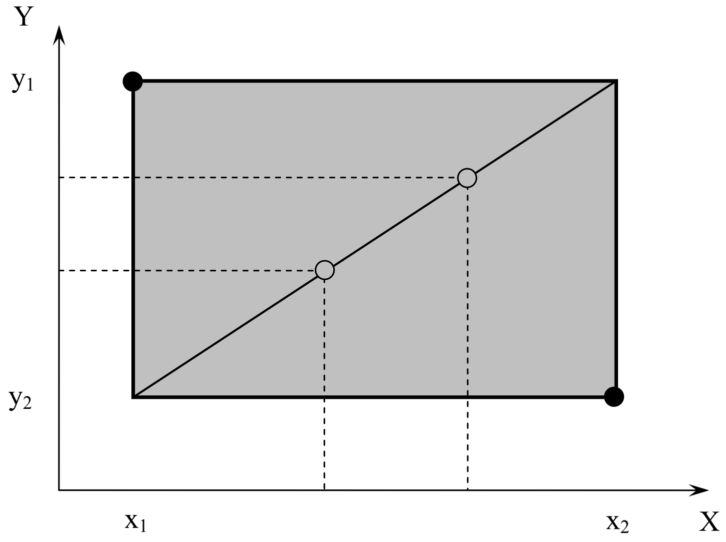

Fig. 2: Comonotonic set is a "thin" set

A support of a random vector X can be obtained by subtracting from $\mathbb{R}^n$ all points which have a zero-probability neighborhood. Roughly speaking, the support is the set of all possible outcomes of X.

Now define comonotonicity of random vectors.

Definition 3.4: A random vector $X = (X_1, X_2, ..., X_n)$ itself and its joint distribution are referred to as *comonotonic* if the random vector X has a comonotonic support.

Therefore, comonotonicity is a very strong dependency structure. If x and y are two elements of the support of X, in other words x and y are possible outcomes of X, then they are ordered componentwise. This explains where the name comonotonic (common monotonic) came from.

The next theorem gives equivalent characterizations for comonotonicity of a random vector, some of them are especially important for future derivation of convex bounds for V.

Theorem 3.1: A random vector $X = (X_1, X_2, ..., X_n)$ is comonotonic (the random variables $X_1, X_2, ..., X_n$ are said to be mutually comonotonic) if and only if any of the following conditions holds:

(1) X has a comonotonic support.

(2) For the n-variate cumulative distribution function we have

$$F_X(x) = \min\{F_{X_1}(x_1), F_{X_2}(x_2), ..., F_{X_n}(x_n)\}. \tag{11}$$

(3) For any random variable U uniformly distributed on (0,1), we have

$$X \stackrel{d}{=} \{F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), ..., F_{X_n}^{-1}(U)\}. \tag{12}$$

(4) There exist a random variable Z and non-decreasing functions $f_i : \mathbb{R} \to \mathbb{R}$ ($i = 1, 2, ..., n$), such that

$$X \stackrel{d}{=} \{f_1(Z), f_2(Z), ..., f_n(Z)\}. \tag{13}$$

In the definitions above, the notation $\stackrel{d}{=}$ is used to indicate that two multivariate random variables are equal in distribution.

Proof (follows the ideas given in Dhaene et al. (2002a)).

(1)$\Longrightarrow$(2): Assume that X has a comonotonic support B. Let $x \in \mathbb{R}^n$ and let A_j be defined by

$$A_j = \{y \in B | y_j \leq x_j\}, \quad j = 1, 2, ..., n.$$

Because of the comonotonicity of B, there exists an i such that $A_i = \bigcap_{j=1}^{n} A_j$. The last assertion can be proved in the following way. It is obvious, that

$$\bigcap_{j=1}^{n} A_j = \{y \in B | y_1 \leq x_1 \cap ... \cap y_n \leq x_n\}.$$

It stands to reason, that $\bigcap_{j=1}^{n} A_j \subseteq A_i$ for all i. The only thing which must be shown is that there exist an $i \in \{1, ..., n\}$ such that

$$A_i \subseteq \bigcap_{j=1}^{n} A_j. \tag{14}$$

Let assume by contradiction that

(1) $A_1 \notin \bigcap_{j=1}^{n} A_j$ and

(2) $A_2 \notin \bigcap_{j=1}^{n} A_j$ and

$\vdots$

(n) $A_n \notin \bigcap_{j=1}^{n} A_j$.

Each of the contradicting assertions implies

(1) $\Longrightarrow \exists\ z^1 \in A_1$ with $z_1^1 \leq x_1,\ z_2^1 > x_2$ or ... or $z_n^1 > x_n$.

(2) $\Longrightarrow \exists\ z^2 \in A_2$ with $z_2^2 \leq x_2,\ z_1^2 > x_1$ or ... or $z_n^2 > x_n$.

$\vdots$

$(n) \Longrightarrow \exists\ z^n \in A_n$ with $z_n^n \leq x_n,\ z_1^n > x_1$ or ... or $z_{n-1}^n > x_{n-1}$.

Let consider the first contradicting assertion, that $A_1 \notin \bigcap_{j=1}^{n} A_j$. This implies that there exist an index $k \in \{2, ..., n\}$ such that $z_k^1 > x_k$. Consequently, on the basis of the first inequalities in assertions (2)-(n) $z_k^1 > x_k \geq z_k^k \Longrightarrow z_k^1 > z_k^k$. Because of comonotonicity it follows $z_m^1 \geq z_m^k$, $m = 1, ..., n$. This implies that for $m = 1$ $x_1 \geq z_1^1 \geq z_1^k$. This leads to $z_1^k \leq x_1$. Therefore, k cannot longer arise as index in (2)-(n).

By analogy consider the second contradicting assertion $A_2 \notin \bigcap_{j=1}^{n} A_j$. This implies there exists an index $l \in \{1, ..., n\} \setminus \{2, k\}$, such that $z_l^2 > x_l$ and $z_l^2 > x_l \geq z_l^l$. Consequently, $z_l^2 > z_l^l$. From comonotonicity it follows $z_m^2 > z_m^l\ \forall\ m = 1, ..., n$. Then taking $m = 2$ one obtains $x_2 \geq z_2^2 \geq z_2^l$, which yields $z_2^l \leq x_2$. Thus, l cannot longer arise as index in (3)-(n).

The repeated application of the described above procedures up to and including finally assertion (n) leads to the proof of assertion (14).

Hence, there exists an index $i \in \{1, ..., n\}$ with

$$\begin{aligned} F_X(x) &= P(\bigcap_{j=1}^{n}\{X_j \leq x_j\}) \\ &= P(X \in \bigcap_{j=1}^{n} A_j) = P(X \in A_i) = P(X_i \leq x_i) \\ &= F_{X_i}(x_i) = \min\{F_{X_1}(x_1), F_{X_2}(x_2), ..., F_{X_n}(x_n)\}. \end{aligned}$$

The last equality follows from $A_i \subset A_j$ so that $F_{X_i}(x_i) \leq F_{X_j}(x_j)$ holds for all values of j.

(2)$\Longrightarrow$(3): Now assume that $F_X(x) = \min\{F_{X_1}(x_1), F_{X_2}(x_2), ..., F_{X_n}(x_n)\}$ for all $x = (x_1, x_2, ..., x_n)$. Then from equation (4) we find that

$$\begin{aligned} P(F_{X_1}^{-1}(U) \leq x_1, ..., F_{X_n}^{-1}(U) \leq x_n) &= P(U \leq F_{X_1}(x_1), ..., U \leq F_{X_n}(x_n)) \\ &= P(U \leq \min_{j=1,...,n}\{F_{X_j}(x_j)\}) \\ &= \min_{j=1,...,n}\{F_{X_j}(x_j)\} = F_X(x). \end{aligned}$$

(3)$\Longrightarrow$(4): straightforward, as the inverse of distribution function is defined as non-decreasing function.

(4)$\Longrightarrow$(1): Assume that there exists a random variable Z with support B, and non-decreasing functions f_i, $(i = 1, 2, ..., n)$, such that

$$X \stackrel{d}{=} \{f_1(Z), f_2(Z), ..., f_n(Z)\}.$$

The set of possible outcomes of X, i.e. support of X, is $\{(f_1(z), f_2(z), ..., f_n(z))|z \in B\}$. This set is obviously comonotonic, which implies that X is indeed comonotonic. ⊠

The assertion (11) in Theorem 3.1 means that in order to find the probability of all the outcomes of n comonotonic risks X_i being less than x_i, $(i = 1, ..., n)$, the probability of the least likely of these n-events must be taken.

It can be seen from condition (12) as well as from condition (13) that comonotonic random variables possess a very strong positive dependence: increasing one of the X_i will lead to an increase of all the other random variables X_j involved. It should be emphasized, that this special random variables provide us with a tool to construct a close upper bound for V. It will be shown later that replicating the dependency copula in the sum (3) by the comonotonic copula $(F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), ..., F_{X_n}^{-1}(U))$ yields an upper bound for V.

The concept of comonotonicity can be explained in terms of Monte Carlo simulation. The risk X and Y are comonotonic if and only if $(X, Y) \stackrel{d}{=} (F_X^{-1}(U), F_Y^{-1}(U))$, for U being any uniformly distributed random variable on (0,1). Hence, in order to simulate comonotonic risks, one needs to generate only one sample of random uniform numbers and insert in F_X^{-1} and F_Y^{-1} to get a sample of pairs (X, Y).

It can be seen from condition (12) that in the special case that all marginal distribution functions F_{X_i} are identical, comonotonicity of X is equivalent to the condition, that $X_1 = X_2 = ... = X_n$ holds with probability one. Taking into account, that the random vectors $(F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), ..., F_{X_n}^{-1}(U))$ and $(F_{X_1}^{-1(\alpha_1)}(U), F_{X_2}^{-1(\alpha_2)}(U), ..., F_{X_n}^{-1(\alpha_n)}(U))$ are equal almost surely, the comonotonicity of X can be also characterized by

$$X \stackrel{d}{=} (F_{X_1}^{-1(\alpha_1)}(U), F_{X_2}^{-1(\alpha_2)}(U), ..., F_{X_n}^{-1(\alpha_n)}(U))$$

for U uniformly distributed on (0,1) and real numbers $\alpha_i \in [0,1]$.

Moreover, if $U \sim$ uniform (0,1), then $1-U \sim$ uniform (0,1). This leads to another equivalent characterization of comonotonic random vector X, i.e.

$$X \stackrel{d}{=} (F_{X_1}^{-1}(1-U), F_{X_2}^{-1}(1-U), ..., F_{X_n}^{-1}(1-U)).$$

Theorem 3.2: A random vector X is comonotonic if and only if (X_i, X_j) is comonotonic for all $i, j \in \{1, 2, .., n\}$, $i \neq j$.

Proof (follows the ideas given in Dhaene et all. (2002a)). "$\Longrightarrow$" straightforward
"$\Longleftarrow$" Consider the set $A \in \mathbb{R}^n$ given by

$$A = \left\{ \left(F_{X_1}^{-1}(p), F_{X_2}^{-1}(p), ..., F_{X_n}^{-1}(p)\right) | 0 < p < 1 \right\}.$$

Its (i, j)-projections are given by

$$A_{i,j} = \left\{ \left(F_{X_i}^{-1}(p), F_{X_j}^{-1}(p)\right) | 0 < p < 1 \right\}.$$

The events $\{X \in A\}$ and $\{(X_i, X_j) \in A_{i,j}$ for all pairs $(i, j)\}$ are equivalent. Pairs (X_i, X_j) are comonotonic by assumption, thus the latter event holds with probability one, i.e. $P((X_i, X_j) \in A_{i,j}) = 1$. Hence $P(X \in A) = 1$. In accordance with Definition 3.3 comonotonic set A is the support of X. Thus in accordance with the definition of comonotonic random vector (see Definition 3.4) X is comonotonic. ⊠

Consequently, the comonotonicity of a random vector is equivalent to pairwise comonotonicity.

The following definition is used later on.

Definition 3.5: Let $X = (X_1, X_2, ..., X_n)$ be a random vector. A random vector with the same marginal distributions and with the comonotonic dependency structure is called *comonotonic counterpart*. It will be denoted by $X^c = (X_1^c, X_2^c, ..., X_n^c)$. Moreover

$$(X_1^c, X_2^c, ..., X_n^c) \stackrel{d}{=} (F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), ..., F_{X_n}^{-1}(U)),$$

where $\stackrel{d}{=}$ means equality in distribution.

3.2 Examples of comonotonicity

The aim of this section is to describe the most widespread situations where comonotonic random variables occur.

In financial context, the most obvious examples of comonotonicity emerge when considering the pay-offs of derivative securities. Such pay-off functions are strongly dependent on the value of the underlying.

Specially, let $S(t)$ be the value of a security at a future time t, $t \geq 0$. Consider a European call option on this security, with expiration date $T \geq 0$ and exercise price K. The pay-off of the call-option at time T is given by $(S(T) - K)_+$. Then the pay-off of a portfolio consisting of the security and the call-option is given by $(S(T), (S(T) - K)_+)$, which is obviously a comonotonic random vector, since the pay-off of the option is a non-decreasing function of the value of the underlying security at the expiration date. Therefore, the holder of the security who buys the call option increases his potential gains. On the other hand, if the holder of the security decides to write the call option, the pay-off of his portfolio at time T is given by $(S(T), -(S(T)-K)_+)$. The pay-off $-(S(T)-K)_+$ is a non-increasing function of $S(T)$. $(S(T), -(S(T)-K)_+)$ is referred to as "counter comonotonic" random vector, if one of the components increases, then the other decreases. Thus writing the call option leads to immediate gain at the expense of reducing the gain of the security.

By analogy, holding the security and buying a European put option leads to a portfolio pay-off $(S(T), (K-S(T))_+)$, which is a counter-comonotonic random vector. If the holder of the security writes a put option than the pay-off of a portfolio is given as $(S(T), -(K-S(T))_+)$, which is obviously a comonotonic random vector.

Comonotonicity is useful instrument for hedging. For example, let $(X_1, X_2, ..., X_n)$ be an insurance portfolio of individual risks X_i, which are not assumed to be mutually independent. In this case, the risks might not be pooled as effectively as expected. There exist the possibility to reduce the aggregate risk of portfolio by financial hedging techniques. The insurer can buy a financial contract with payments Y, such that Y and $X_1+X_2+...+X_n$ are comonotonic. Then the increase of aggregate loss $X_1+X_2+...+X_n$ will be compensated by the increase of the payment of financial contract. On the other hand, the insurer could also sell a financial contract with obligations for him, which are negatively correlated with the aggregate loss. Then the increase of aggregate insurance loss will be compensated by the decrease of the obligations, related to the financial contract.

Another straightforward example of comonotonicity occuring in insurance environment is the present value of annuities. Consider a life annuity on a life aged x, which has to pay a unit amount at the end of each year, provided that he/she is still alive. The future lifetime of a life aged x is denoted by T. It is a non-negative continuous random variable with a probability distribution function $G(t) = P(T \leq t)$. The function $G(t)$ represents the probability that the person will die within t years for any fixed t. In accordance with the actuarial community's system of notations the distribution function of T is denoted by ${}_tq_x$, i.e. ${}_tq_x = G(t)$.

Let ω be the ultimate age of the life table, i.e. $\omega - x$ is the first remaining lifetime of a life aged x for which ${}_{\omega-x}q_x = 1$ fulfills. Under assumption that discounting is performed with deterministic interest rate r, the present value S of the future payment is equal to

$$S = \sum_{i=1}^{\lceil \omega - x \rceil - 1} X_i, \tag{15}$$

where $\lceil \cdot \rceil$ is the number of completed years of $\omega - x$, and where X_i are given by

$$X_i = \left(\frac{1}{1+r}\right)^i I_{(T>i)}.$$

Clearly, that the present value of annuity (15) is a comonotonic sum, as its components X_i are non-decreasing functions of the future life time T. From the same reason the payment vector $X = (X_1, X_2, ..., X_n)$ is also comonotonic.

4 Convex bounds for sums of random variables

4.1 General results

In the modern actuarial literature several proofs have been given for the fact that the sum of the components $X_1 + X_2 + ... + X_n$ is the riskiest (largest in the convex order), if the joint distribution of the random vector $(X_1, X_2, ..., X_n)$ is comonotonic, i.e. risks $X_1, X_2, ..., X_n$ are comonotonous. The comonotonous distribution is that of the vector $(F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), ..., F_{X_n}^{-1}(U))$ (see Theorem 3.1). The components of this random vector are maximally dependent, all being non-decreasing functions of the same random variable. Technically speaking, replicating in the Formula (3) the copula $X_1, X_2, ..., X_n$ with X_k given by

$$X_k = \widehat{\alpha}_k e^{\widehat{Z}(k)}, \quad k = 1, ..., n$$

by comonotonic copula yields an upper bound for V in the convex order.

Independent proofs of this central result have appeared is several papers. Proofs of it involving the general concept of supermodularity are contained in Goovaerts and Dhaene (1999), Goovaerts and Redant (1999) and Bäuerle and Müller (1998). Proofs for the special case $n = 2$ and for individual life models (two point distributions) can be found in Dhaene and Goovaerts (1995, 1996). Goovaerts et al. (2000) consider this theorem only in the case of continuous random variables. Dhaene et al. (2000) prove a slightly more general result. It should be also noted that a geometric proof based on the properties of the comonotonic support is given in Kaas et al. (2001). A general proof using an extension of the notion of inverse distribution function will be given below. It follows the ideas of Kaas et al. (2000).

Theorem 4.1: (Convex upper bound for sums of random variables)
Let U be a uniformly (0,1) distributed random variable. For any random vector $X = (X_1, X_2, ..., X_n)$ with marginal cumulative distribution functions $F_{X_1}, F_{X_2}, ..., F_{X_n}$ we have that

$$\sum_{i=1}^{n} X_i \leq_{cx} \sum_{i=1}^{n} F_{X_i}^{-1}(U).$$

The following two results, which can be easily proved on the basis of Remark 1.2 and Definition 1.2 of the α-inverse of the distribution function, are important for the proof of Theorem 4.1.

Remark 4.1: If $V^c \stackrel{d}{=} F_{X_1}^{-1}(U) + F_{X_2}^{-1}(U) + ... + F_{X_n}^{-1}(U)$ (d means equality in distribution), then the inverse cumulative distribution function of this sum of comonotonic random variables can be calculated as

$$F_{V^c}^{-1}(p) = \sum_{i=1}^{n} F_{X_i}^{-1}(p), \quad p \in (0, 1). \tag{16}$$

Analogously it holds

$$F_{V^c}^{-1(\alpha)}(p) = \sum_{i=1}^{n} F_{X_i}^{-1(\alpha)}(p), \quad p \in (0, 1). \tag{17}$$

Proof of the Theorem 4.1. From the reason of convenience make the following notation $V \stackrel{d}{=} X_1 + X_2 + ... + X_n$. Obviously $E[V] = E[V^c]$. Consequently, in accordance with

Definition 2.3 to prove convex order inequality the stop-loss inequality $X \leq_{sl} Y$ must be shown. For this aim consider an arbitrary number d such that $0 < F_{V^c}(d) < 1$. Take $\alpha \in [0,1]$ such that $F_{V^c}^{-1(\alpha)}(F_{V^c}(d)) = d$. This relation is possible on the basis of the Definition 1.2 of the α-inverse of cumulative distribution function and Remark 1.1. Thus

$$E[V-d]_+ = E[V - F_{V^c}^{-1(\alpha)}(F_{V^c}(d))]_+.$$

The α-inverse of a cumulative distribution function can be easily computed according to Formula (17). Hence

$$\begin{aligned} &E[V - F_{V^c}^{-1(\alpha)}(F_{V^c}(d))]_+ \\ &= E\left[\sum_{i=1}^{n}(X_i - F_{X_i}^{-1(\alpha)}(F_{V^c}(d)))\right]_+ \leq \sum_{i=1}^{n} E[X_i - F_{X_i}^{-1(\alpha)}(F_{V^c}(d))]_+. \end{aligned}$$

As the next step the stop-loss premium at retention d of V^c will be calculated.

$$\begin{aligned} E[V^c - d]_+ &= E[F_{V^c}^{-1}(U) - d]_+ \\ &= \int_0^1 (F_{V^c}^{-1}(p) - d)_+ dp \\ &= \int_0^1 (F_{V^c}^{-1}(p) - F_{V^c}^{-1(\alpha)}(F_{V^c}(d)))_+ dp \\ &= \int_{F_{V^c}(d)}^1 (F_{V^c}^{-1}(p) - F_{V^c}^{-1(\alpha)}(F_{V^c}(d))) dp. \end{aligned}$$

Applying relations (16) and (17) for the calculation of inverse and α-inverse cumulative distribution function of comonotonous random variables respectively one can derive

$$E[V^c - d]_+ = \sum_{i=1}^{n} \int_{F_{V^c}(d)}^1 (F_{X_i}^{-1}(p) - F_{X_i}^{-1(\alpha)}(F_{V^c}(d))) dp.$$

For any $p \in (F_{V^c}(d); F_{X_i}(F_{X_i}^{-1(\alpha)}(F_{V^c}(d))))$ it holds

$$F_{X_i}^{-1}(p) = F_{X_i}^{-1(\alpha)}(F_{V^c}(d)).$$

Consequently

$$\begin{aligned} E[V^c - d]_+ &= \sum_{i=1}^{n} \int_{F_{X_i}(F_{X_i}^{-1(\alpha)}(F_{V^c}(d)))}^1 (F_{X_i}^{-1}(p) - F_{X_i}^{-1(\alpha)}(F_{V^c}(d)))) dp \\ &= \sum_{i=1}^{n} \int_0^1 (F_{X_i}^{-1}(p) - F_{X_i}^{-1(\alpha)}(F_{V^c}(d)))_+ dp \\ &= \sum_{i=1}^{n} E[F_{X_i}^{-1}(U) - F_{X_i}^{-1(\alpha)}(F_{V^c}(d))]_+ \\ &= \sum_{i=1}^{n} E[X_i - F_{X_i}^{-1(\alpha)}(F_{V^c}(d))]_+. \end{aligned}$$

Thus $E[V-d]_+ \leq E[V^c - d]_+$ holds for all d with $0 < F_{V^c}(d) < 1$ as it was assumed.

Now the same inequality must be checked for retention d with $F_{V^c}(d) = 0$ and $F_{V^c}(d) = 1$. As the stop-loss transform is a continuous non-increasing function of d, the following two inequalities are true:

$$E[V - F_{V^c}^{-1+}(0)]_+ \leq E[V^c - F_{V^c}^{-1+}(0)]_+,$$
$$E[V - F_{V^c}^{-1}(1)]_+ \leq E[V^c - F_{V^c}^{-1}(1)]_+.$$

Consider the situation when $d < F_{V^c}^{-1+}(0)$. It holds

$$E[V^c - F_{V^c}^{-1+}(0)]_+ = E[V^c - F_{V^c}^{-1+}(0)] = E[V - F_{V^c}^{-1+}(0)].$$

Obviously

$$E[V - F_{V^c}^{-1+}(0)]_+ \geq E[V - F_{V^c}^{-1+}(0)].$$

Consequently

$$E[V - F_{V^c}^{-1+}(0)] = E[V - F_{V^c}^{-1+}(0)]_+.$$

From the last relation it follows

$$P(V < F_{V^c}^{-1+}(0)) = 0,$$

and therefore it holds

$$P(V < d) = 0$$

under initial assumption that $d < F_{V^c}^{-1+}(0)$. Hence

$$E[V - d]_+ = E[V - d] = E[V^c - d] = E[V^c - d]_+,$$

which proves that relation $E[V-d]_+ \leq E[V^c-d]_+$ holds with equality sign for $d < F_{V^c}^{-1+}(0)$.

Now assume that $d > F_{V^c}^{-1}(1)$. Hence

$$E[V - d]_+ \leq E[V - F_{V^c}^{-1}(1)]_+ \leq E[V^c - F_{V^c}^{-1}(1)]_+ = 0 = E[V^c - d]_+.$$

This proves that $E[V - d]_+ \leq E[V^c - d]_+$ holds for retention d with $d > F_{V^c}^{-1}(1)$. Thus $E[V - d]_+ \leq E[V^c - d]_+$ holds for all values of d. Last inequality together with equality of means implies that $V \leq_{cx} V^c$. ⊠

If in addition to the upper bound a lower bound can be found as well, this provides one with a measure of the reliability of the upper bound. Technically speaking, applying Jensen's inequality for conditional expectations,

$$E[f(V)|\Lambda] \geq f(E[V|\Lambda])$$

where $f : \mathbb{R} \longrightarrow \mathbb{R}$ is a convex function, yields a convex lower bound for V.

Theorem 4.2: (Convex lower bound for sums of random variables)
For any random vector $X = (X_1, X_2, ..., X_n)$ and any random variable Λ, which is assumed to be a function of the random vector X, we have

$$\sum_{i=1}^{n} E[X_i|\Lambda] \leq_{cx} \sum_{i=1}^{n} X_i.$$

Proof. To prove the convex order inequality it must be shown in accordance with Definition 2.3 of convex order precedence, that $E[\sum_{i=1}^n E[X_i|\Lambda] - d]_+ \leq E[\sum_{i=1}^n X_i - d]_+$ for all $d \in \mathbb{R}$ and, additionally, $E[\sum_{i=1}^n E[X_i|\Lambda]] = E[\sum_{i=1}^n X_i]$. The inequality of stop-loss premiums can be proved applying Jensen's inequality, since $(\cdot)_+ = \max(\cdot, 0)$ is a convex function. More precise

$$\begin{aligned} E[\sum_{i=1}^n X_i - d]_+ &= E[E[(\sum_{i=1}^n X_i - d)_+|\Lambda]] \geq E[E[\sum_{i=1}^n X_i|\Lambda] - d]_+ \\ &= E[\sum_{i=1}^n E(X_i|\Lambda) - d]_+. \end{aligned}$$

The equality of means can be easily shown in the following way

$$E[\sum_{i=1}^n X_i] = E[E(\sum_{i=1}^n X_i|\Lambda)] = E[\sum_{i=1}^n E(X_i|\Lambda)]. \quad \boxtimes$$

It should be remarked, that slightly different proof of Theorem 4.2, based on the equivalent definition of convex order in terms of utility functions (see Lemma 2.1), is contained in Kaas et al. (2000). The technique for deriving a lower bound is also considered (for some special cases) by Vyncke et al. (2001). The idea of this technique stems from mathematical physics.

Two extreme cases are possible when deriving a lower bound for $V = X_1 + X_2 + ... + X_n$ by means of Theorem 4.2. The first case is that $\Lambda = V$, or, saying it in words, Λ and V are in one-to-one relation. In this case the lower bound for V is nothing else as V itself. The second case arises when V and Λ are mutually independent. Then the lower bound turns to be $E[V]$.

From the reason of convenience the following notations will be used throughout this thesis

$$V^c = \sum_{i=1}^n F_{X_i}^{-1}(U) \tag{18}$$

and

$$V^l = \sum_{i=1}^n E[X_i|\Lambda].$$

Components of the sum V^c are maximally dependent, all being non-decreasing functions of the same uniform (0,1) random variable U. Thus V^c is a comonotonic sum, which will be referred to as *comonotonic upper bound* in the context of this thesis. Clearly, that nothing can be said about comonotonicity of the sum V^l from the general form of its components. However, if the conditioning random variable Λ is chosen in such a way that all random variables $E[X_i|\Lambda]$ are non-decreasing functions of Λ (or all are non-increasing functions of Λ), then V^l is a sum of n comonotonous random variables and, thus, V^l will be referred to as *comonotonic lower bound.*

Unification of the general results of Theorems 4.1 and 4.2 leads to a central theoretical result of this thesis, which, under assumption of the notations given above, can be written as

$$V^l \leq_{cx} V \leq_{cx} V^c.$$

The comonotonic upper bound V^c changes the original copula, but keeps the marginal distributions unchanged, since

$$P[F_{X_i}^{-1}(U) \leq x] = P[U \leq F_{X_i}(x)] = F_{X_i}(x).$$

In its turn, the comonotonic lower bound V^l changes both the copula and the marginal distributions involved, since $E[E[X_i|\Lambda]] = E[X_i]$ always holds according to the property of conditional expectation, but $Var[E[X_i|\Lambda]] < Var[X_i]$ unless $E[Var[X_i|\Lambda]] = 0$ which means in principle that X_i, for a given $\Lambda = \lambda$, is a constant for each λ.

It should be noted, that the united results of Theorems 4.1 and 4.2 can be extended to the case of determining the approximations for the distribution function of a scalar product of two mutually independent random vectors given by

$$V = \sum_{i=1}^{n} Y_i X_i.$$

In applications, Y_i can be future random (not deterministic as in current situation) payments due at time $t_0, t_1, ..., t_{n-1}$ and X_i are random discount factors. The technique of constructing convex bounds in this special case can be found in Ahcan et al. (2006).

4.2 Closed-form expressions for valuing convex bounds

To prove the central result of this section the following lemma is important.

Lemma 4.1: Consider a random variable X that is lognormally distributed, i.e. $\ln X \sim N(\mu, \sigma^2)$, then the inverse distribution function of X can be calculated as

$$F_X^{-1}(p) = e^{\mu + \sigma\Phi^{-1}(p)}.$$

Proof. Let Y be a normally $(0, 1)$ distributed random variable. Applying Remark 1.2 (assertion 1) with $g(x) = \exp(\mu + \sigma x)$ one can immediately derive that

$$F_X^{-1}(p) = F_{g(Y)}^{-1}(p) = g(\Phi^{-1}(p)) = \exp(\mu + \sigma\Phi^{-1}(p)),$$

which proves the stated above result. ⊠

Theorem 4.3: (explicit formulas for the convex lower and upper bounds)

Let the random variable V be given by

$$V = \sum_{i=1}^{n} \widehat{\alpha}_i \exp(\widehat{Z}_1 + \widehat{Z}_2 + ... + \widehat{Z}_i) = \sum_{i=1}^{n} \widehat{\alpha}_i e^{\widehat{Z}(i)},$$

where $\widehat{\alpha}_i$ are non-negative real numbers, the random vector $(\widehat{Z}_1, \widehat{Z}_2, ..., \widehat{Z}_n)$ has a multivariate normal distribution and $\widehat{Z}(i)$ are defined as

$$\widehat{Z}(i) = \sum_{k=1}^{i} \widehat{Z}_k. \tag{19}$$

Consider the conditioning random variable given by

$$\Lambda = \sum_{i=1}^{n} \beta_i \widehat{Z}_i. \tag{20}$$

Then the lower bound V^l and the upper bound V^c can be calculated as

$$V^l = \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\frac{1}{2}(1-r_i^2)\sigma^2_{\widehat{Z}(i)}+r_i\sigma_{\widehat{Z}(i)}\Phi^{-1}(U)} \tag{21}$$

and

$$V^c = \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\sigma_{\widehat{Z}(i)}\Phi^{-1}(U)} \tag{22}$$

respectively. Here U is a uniform (0,1) random variable, and r_i are the correlation coefficients between Λ and $\widehat{Z}(i)$ given by

$$r_i = \frac{cov[\widehat{Z}(i), \Lambda]}{\sigma_{\widehat{Z}(i)}\sigma_\Lambda}.$$

Proof (follows the ideas of Dhaene et al. (2002b)). The random vector $(\widehat{Z}_1, \widehat{Z}_2, ..., \widehat{Z}_n)$ has a multivariate normal distribution. Thus, in accordance with Definition 1.4, every linear combination of its components has a univariate normal distribution. So Λ given by (20) as well as $\widehat{Z}(i)$ given by (19) are normally distributed. Moreover, $(\widehat{Z}(i), \Lambda)$ has a bivariate normal distribution. Conditionally given $\Lambda = \lambda$, $\widehat{Z}(i)$ has a univariate normal distribution with mean and variance given by (see Shiryaev (1996), p.238):

$$\mu_i := E[\widehat{Z}(i)|\Lambda = \lambda] = E[\widehat{Z}(i)] + r_i \frac{\sigma_{\widehat{Z}(i)}}{\sigma_\Lambda}(\lambda - E[\Lambda])$$

and

$$\sigma_i^2 := Var[\widehat{Z}(i)|\Lambda = \lambda] = \sigma^2_{\widehat{Z}(i)}(1 - r_i^2).$$

Consequently, given $\Lambda = \lambda$, the random variable

$$\overline{X_i} = e^{\widehat{Z}(i)}$$

is lognormally distributed with parameters μ_i and σ_i^2. The expected value of this lognormally distributed random variable can be easily calculated as

$$E[\overline{X_i}|\Lambda = \lambda] = e^{\mu_i + \frac{1}{2}\sigma_i^2}.$$

From $X_i = \widehat{\alpha}_i e^{\widehat{Z}(i)} = \widehat{\alpha}_i \overline{X_i}$ immediately follows that

$$E[X_i|\Lambda = \lambda] = \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+r_i\sigma_{\widehat{Z}(i)}\frac{\lambda - E[\Lambda]}{\sigma_\Lambda}+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}(1-r_i^2)}.$$

Consequently it holds

$$E[X_i|\Lambda] = \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+r_i\sigma_{\widehat{Z}(i)}\Phi^{-1}(U)+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}(1-r_i^2)},$$

taking into account, that $\frac{\Lambda - E[\Lambda]}{\sigma_\Lambda}$ is $N(0,1)$ distributed and therefore $U \equiv \Phi\left(\frac{\Lambda - E[\Lambda]}{\sigma_\Lambda}\right) \sim U[0,1]$. Hence, the lower bound $V^l = \sum_{i=1}^n E[X_i|\Lambda]$ is given by (21).

Taking into account that $\overline{X_i} = e^{\widehat{Z}(i)}$ is lognormally distributed with parameters $E[\widehat{Z}(i)]$ and $\sigma^2_{\widehat{Z}(i)}$, and in view of the Remark 1.2 (assertion 1) and Lemma 4.1 it can be easily derived that

$$F^{-1}_{X_i}(p) = F^{-1}_{\widehat{\alpha}_i\overline{X_i}}(p) = \widehat{\alpha}_i F^{-1}_{\overline{X_i}}(p) = \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\sigma_{\widehat{Z}(i)}\Phi^{-1}(p)}.$$

From the last expression it follows that $V^c = \sum_{i=1}^n F_{X_i}^{-1}(U)$ is given by (22). ⊠

In accordance with Theorem 3.1 (assertion 3), V^c is a comonotonic sum, as the components of this sum are non-decreasing functions of the same random variable U. Provided that all coefficients r_i are positive, the terms of V^l are also non-decreasing functions of the same random variable U. Thus, in this case V^l is a comonotonic sum too.

Finally, it should be noted that the expected values of the random variables V, V^c and V^l are all equal

$$E(V) = E(V^l) = E(V^c) = \sum_{i=1}^n \widehat{\alpha}_i e^{E[\widehat{Z}(i)] + \frac{1}{2}\sigma^2_{\widehat{Z}(i)}}. \tag{23}$$

To be more precise,

$$E(V) = E\left(\sum_{i=1}^n \widehat{\alpha}_i e^{\widehat{Z}(i)}\right) = \sum_{i=1}^n \widehat{\alpha}_i E e^{\widehat{Z}(i)}.$$

$\widehat{Z}(i)$ is a normally distributed random variable with parameters $E[\widehat{Z}(i)]$ and $\sigma^2_{\widehat{Z}(i)}$. Thus

$$E(V) = \sum_{i=1}^n \widehat{\alpha}_i E e^{\widehat{Z}(i)} = \sum_{i=1}^n \widehat{\alpha}_i e^{E[\widehat{Z}(i)] + \frac{1}{2}\sigma^2_{\widehat{Z}(i)}}.$$

By analogy

$$\begin{aligned} E(V^l) &= E\left(\sum_{i=1}^n \widehat{\alpha}_i e^{E[\widehat{Z}(i)] + \frac{1}{2}(1-r_i^2)\sigma^2_{\widehat{Z}(i)} + r_i \sigma_{\widehat{Z}(i)} \Phi^{-1}(U)}\right) \\ &= \sum_{i=1}^n \widehat{\alpha}_i E e^{E[\widehat{Z}(i)] + \frac{1}{2}(1-r_i^2)\sigma^2_{\widehat{Z}(i)} + r_i \sigma_{\widehat{Z}(i)} \Phi^{-1}(U)}. \end{aligned}$$

As $\Phi^{-1}(U)$ is a $N(0,1)$ distributed random variable, the exponent expression in the last formula is a normally distributed random variable with the mean $E[\widehat{Z}(i)] + \frac{1}{2}(1-r_i^2)\sigma^2_{\widehat{Z}(i)}$ and variance $r_i^2\sigma^2_{\widehat{Z}(i)}$. Hence

$$E(V^l) = \sum_{i=1}^n \widehat{\alpha}_i e^{E[\widehat{Z}(i)] + \frac{1}{2}(1-r_i^2)\sigma^2_{\widehat{Z}(i)} + \frac{1}{2}r_i^2\sigma^2_{\widehat{Z}(i)}} = \sum_{i=1}^n \widehat{\alpha}_i e^{E[\widehat{Z}(i)] + \frac{1}{2}\sigma^2_{\widehat{Z}(i)}}.$$

The expected value of V^c can be derived by analogy.

It can be easily shown that the variances of V, V^l and V^c are given by

$$\begin{aligned} Var(V) &= \sum_{i=1}^n \sum_{j=1}^n \widehat{\alpha}_i \widehat{\alpha}_j e^{E[\widehat{Z}(i)] + E[\widehat{Z}(j)] + \frac{1}{2}(\sigma^2_{\widehat{Z}(i)} + \sigma^2_{\widehat{Z}(j)})} (e^{cov(\widehat{Z}(i), \widehat{Z}(j))} - 1), \\ Var(V^l) &= \sum_{i=1}^n \sum_{j=1}^n \widehat{\alpha}_i \widehat{\alpha}_j e^{E[\widehat{Z}(i)] + E[\widehat{Z}(j)] + \frac{1}{2}(\sigma^2_{\widehat{Z}(i)} + \sigma^2_{\widehat{Z}(j)})} (e^{r_i r_j \sigma_{\widehat{Z}(i)} \sigma_{\widehat{Z}(j)}} - 1), \qquad (24) \\ Var(V^c) &= \sum_{i=1}^n \sum_{j=1}^n \widehat{\alpha}_i \widehat{\alpha}_j e^{E[\widehat{Z}(i)] + E[\widehat{Z}(j)] + \frac{1}{2}(\sigma^2_{\widehat{Z}(i)} + \sigma^2_{\widehat{Z}(j)})} (e^{\sigma_{\widehat{Z}(i)} \sigma_{\widehat{Z}(j)}} - 1). \end{aligned}$$

For example, the variance of V^l can be calculated in the following way

$$Var(V^l) = E(V^l)^2 - (E(V^l))^2. \tag{25}$$

$$\begin{aligned}
E(V^l)^2 &= E\sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(1-r_i^2)\sigma^2_{\widehat{Z}(i)}+\frac{1}{2}(1-r_j^2)\sigma^2_{\widehat{Z}(j)}+(r_i\sigma_{\widehat{Z}(i)}+r_j\sigma_{\widehat{Z}(j)})\Phi^{-1}(U)}\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j Ee^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(1-r_i^2)\sigma^2_{\widehat{Z}(i)}+\frac{1}{2}(1-r_j^2)\sigma^2_{\widehat{Z}(j)}+(r_i\sigma_{\widehat{Z}(i)}+r_j\sigma_{\widehat{Z}(j)})\Phi^{-1}(U)}\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(1-r_i^2)\sigma^2_{\widehat{Z}(i)}+\frac{1}{2}(1-r_j^2)\sigma^2_{\widehat{Z}(j)}+\frac{1}{2}(r_i^2\sigma^2_{\widehat{Z}(i)}+r_j^2\sigma^2_{\widehat{Z}(j)})+r_ir_j\sigma_{\widehat{Z}(i)}\sigma_{\widehat{Z}(j)}}\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)})+r_ir_j\sigma_{\widehat{Z}(i)}\sigma_{\widehat{Z}(j)}}.
\end{aligned}$$

From (23) follows that

$$(E(V^l))^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)})}.$$

Consequently, applying Formula (25) the variance of the lower bound can be calculated as

$$Var(V^l) = \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)})}(e^{r_ir_j\sigma_{\widehat{Z}(i)}\sigma_{\widehat{Z}(j)}} - 1). \tag{26}$$

Obviously, that from the chain of convex order inequalities

$$V^l \leq_{cx} V \leq_{cx} V^c$$

immediately follows (see Remark 2.2) that

$$Var(V^l) \leq Var(V) \leq Var(V^c).$$

In particular, the following relation holds true between the variances of the random variables V^l and V:

$$Var(V) = Var(V^l) + E[Var[V|\Lambda]]. \tag{27}$$

This follows on the one hand from the general relation

$$Var(X) = Var(E[X|\Lambda]) + E[Var(X|\Lambda)],$$

which is true for all random variables X for which the corresponding values exist. On the other hand it can be seen directly from the following derivations:

$$E[V^2|\Lambda] = \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j E[e^{\widehat{Z}(i)+\widehat{Z}(j)}|\Lambda].$$

As $(\widehat{Z}(i)+\widehat{Z}(j),\Lambda)$ has bivariate normal distribution, then, conditionally given Λ, $\widehat{Z}(i)+\widehat{Z}(j)$ is normally distributed with mean and variance given by:

$$\begin{aligned}\mu_{ij} &:= E[\widehat{Z}(i)+\widehat{Z}(j)|\Lambda=\lambda]\\ &= E[\widehat{Z}(i)+\widehat{Z}(j)]+r(\widehat{Z}(i)+\widehat{Z}(j),\Lambda)\frac{\sigma_{\widehat{Z}(i)+\widehat{Z}(j)}}{\sigma_\Lambda}(\lambda-E[\Lambda]),\end{aligned}$$

$$\sigma_{ij}^2 := Var[\widehat{Z}(i)+\widehat{Z}(j)|\Lambda=\lambda]=\sigma^2_{\widehat{Z}(i)+\widehat{Z}(j)}(1-r(\widehat{Z}(i)+\widehat{Z}(j),\Lambda)^2),$$

where $r(\cdot,\cdot)$ denotes the corresponding correlation coefficient. Consequently, conditionally given $\Lambda=\lambda$, $e^{\widehat{Z}(i)+\widehat{Z}(j)}$ is lognormally distributed with parameters μ_{ij} and σ_{ij}^2. Thus

$$E[e^{\widehat{Z}(i)+\widehat{Z}(j)}|\Lambda=\lambda]=e^{\mu_{ij}+\frac{1}{2}\sigma_{ij}^2}.$$

Therefore

$$E[V^2|\Lambda]=\sum_{i=1}^n\sum_{j=1}^n\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)+\widehat{Z}(j)]+r(\widehat{Z}(i)+\widehat{Z}(j),\Lambda)\sigma_{\widehat{Z}(i)+\widehat{Z}(j)}\Phi^{-1}(U)+\frac{1}{2}\sigma^2_{\widehat{Z}(i)+\widehat{Z}(j)}(1-r(\widehat{Z}(i)+\widehat{Z}(j),\Lambda)^2)}.$$

Further

$$E[V|\Lambda]=E\left[\sum_{i=1}^n\widehat{\alpha}_i e^{\widehat{Z}(i)}|\Lambda\right]=\sum_{i=1}^n\widehat{\alpha}_i E[e^{\widehat{Z}(i)}|\Lambda].$$

Conditionally given Λ, $\widehat{Z}(i)$ is normally distributed with parameters

$$\mu_i := E[\widehat{Z}(i)|\Lambda=\lambda]=E[\widehat{Z}(i)]+r(\widehat{Z}(i),\Lambda)\frac{\sigma_{\widehat{Z}(i)}}{\sigma_\Lambda}(\lambda-E[\Lambda])$$

and

$$\sigma_i^2 := Var[\widehat{Z}(i)|\Lambda=\lambda]=\sigma^2_{\widehat{Z}(i)}(1-r(\widehat{Z}(i),\Lambda)^2).$$

Consequently

$$E[V|\Lambda=\lambda]=\sum_{i=1}^n\widehat{\alpha}_i e^{\mu_i+\frac{1}{2}\sigma_i^2}$$

and

$$E[V|\Lambda]=\sum_{i=1}^n\widehat{\alpha}_i e^{E[\widehat{Z}(i)]+r_i\sigma_{\widehat{Z}(i)}\Phi^{-1}(U)+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}(1-r_i^2)}.$$

Thus

$$(E[V|\Lambda])^2=\sum_{i=1}^n\sum_{j=1}^n\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+(r_i\sigma_{\widehat{Z}(i)}+r_j\sigma_{\widehat{Z}(j)})\Phi^{-1}(U)+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}(1-r_i^2)+\frac{1}{2}\sigma^2_{\widehat{Z}(j)}(1-r_j^2)}.$$

Hence

$$E\left[Var[V|\Lambda]\right] = E\left[E[V^2|\Lambda]\right] - E\left[(E[V|\Lambda])^2\right]$$
$$= E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)+\widehat{Z}(j)]+r(\widehat{Z}(i)+\widehat{Z}(j),\Lambda)\sigma_{\widehat{Z}(i)+\widehat{Z}(j)}\Phi^{-1}(U)+\frac{1}{2}\sigma^2_{\widehat{Z}(i)+\widehat{Z}(j)}(1-r(\widehat{Z}(i)+\widehat{Z}(j),\Lambda)^2)}\right]$$
$$-E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+(r_i\sigma_{\widehat{Z}(i)}+r_j\sigma_{\widehat{Z}(j)})\Phi^{-1}(U)+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}(1-r_i^2)+\frac{1}{2}\sigma^2_{\widehat{Z}(j)}(1-r_j^2)}\right].$$

$$\begin{aligned} E\left[E[V^2|\Lambda]\right] &= \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)}+2cov(\widehat{Z}(i),\widehat{Z}(j)))(1-r(\widehat{Z}(i)+\widehat{Z}(j),\Lambda)^2)} \\ &\times\ e^{\frac{1}{2}(r(\widehat{Z}(i)+\widehat{Z}(j),\Lambda))^2(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)}+2cov(\widehat{Z}(i),\widehat{Z}(j)))} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)})+cov(\widehat{Z}(i),\widehat{Z}(j))}. \end{aligned}$$

$$E\left[(E[V|\Lambda])^2\right] =$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}(1-r_i^2)+\frac{1}{2}\sigma^2_{\widehat{Z}(j)}(1-r_j^2)+\frac{1}{2}r_i^2\sigma^2_{\widehat{Z}(i)}+\frac{1}{2}r_j^2\sigma^2_{\widehat{Z}(j)}+r_i\sigma_{\widehat{Z}(i)}r_j\sigma_{\widehat{Z}(j)}}$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)})+r_i\sigma_{\widehat{Z}(i)}r_j\sigma_{\widehat{Z}(j)}}.$$

Therefore

$$E\left[Var[V|\Lambda]\right] = \sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)})}\left[e^{cov(\widehat{Z}(i),\widehat{Z}(j))} - e^{r_i\sigma_{\widehat{Z}(i)}r_j\sigma_{\widehat{Z}(j)}}\right]. \quad (28)$$

Summing (28) and (24) proves Formula (27) by direct calculation.

4.3 Choice of the conditioning random variable

As was already mentioned in Section 4.2, the comonotonicity of the convex lower bound is strongly dependent on the special choice of the conditioning random variable Λ. To be more precise, the conditioning random variable Λ must be chosen in such a way that correlation coefficients r_i are positive (see Section 6.1 for details). Recall, that if $X \leq_{cx} Y$, then $Var[X] \leq Var[Y]$ (see Remark 2.2). It becomes intuitively clear, that if one wants to replace V by the less convex V^l, then the best approximation arise when the variance of V^l is as close as possible to the variance of V. Hence one has to choose Λ such that the ratio $z = \frac{Var(V^l)}{Var(V)}$ is as close as possible to one. The first possible way to choose Λ can be realized by means of numerical procedures for optimizing z. Unfortunately, this technique has an essential drawback: quantiles and Conditional Tail Expectations of V^l can not be easily obtained for such a choice of Λ. Thus, there arise a necessity to obtain a ready-to-use approximations that can be easily implemented and applied by all kinds of users.

Two slightly different versions of conditioning random variable are proposed in literature. Consider, as was agreed, a random variable Λ, which is given as a linear combination of the random variables $\widehat{Z}_i$:

$$\Lambda = \sum_{i=1}^{n} \beta_i \widehat{Z}_i,$$

for particular choices of the coefficients β_i.

Kaas et al. (2000) (see also Dhaene et al. (2000b)) propose the following choice for the parameters β_i:

$$\beta_i = \sum_{j=i}^{n} \widehat{\alpha}_j e^{E[\widehat{Z}(j)]}, \qquad i = 1, ..., n,$$

where $\widehat{Z}(j) = \sum_{k=1}^{j} \widehat{Z}_k$. This special choice of β_i, which will be referred to as *lower bound approach*, makes Λ a linear transformation of a first order approximation to V. This becomes obvious from the following derivation:

$$\begin{aligned}
V &= \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+(\widehat{Z}(i)-E[\widehat{Z}(i)])} \\
&= \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\sum_{j=1}^{i}(\widehat{Z}_j-E[\widehat{Z}_j])} \\
&= \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]} \left[1 + \sum_{j=1}^{i}(\widehat{Z}_j - E[\widehat{Z}_j]) + o(\widehat{Z}_j - E[\widehat{Z}_j])\right] \\
&= C + \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]} \sum_{j=1}^{i} \widehat{Z}_j + o(\sum_{j=1}^{i}(\widehat{Z}_j - E[\widehat{Z}_j])) \\
&= C + \sum_{j=1}^{n} \widehat{Z}_j \sum_{i=j}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]} + o(\sum_{j=1}^{i}(\widehat{Z}_j - E[\widehat{Z}_j])) \\
&= C + \sum_{j=1}^{n} \widehat{Z}_j \beta_j + o(\sum_{j=1}^{i}(\widehat{Z}_j - E[\widehat{Z}_j])) = C + \Lambda + o(\sum_{j=1}^{i}(\widehat{Z}_j - E[\widehat{Z}_j])),
\end{aligned}$$

where C is the constant given by $C = \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]}(1-E[\widehat{Z}(i)])$. Hence V^l will be "close" to V, provided $\widehat{Z}_j - E[\widehat{Z}_j]$ are sufficiently small or, equivalently, $\sigma^2_{\widehat{Z}_j}$ are sufficiently small. For this choice of Λ, $E[Var[V|\Lambda]]$ is "small", as $E[Var[V|\Lambda]] \approx E[Var[C+\Lambda|\Lambda]] = 0$. Since $Var(V) = E[Var[V|\Lambda]] + Var(V^l)$ this implies that $z = \frac{Var(V^l)}{Var(V)}$ is expected to tend to one.

The second approach for the conditioning random variable Λ, which is slightly different from the first approach, is proposed by Vanduffel et al. (2005a). They choose parameters β_i equal to

$$\beta_i = \sum_{j=i}^{n} \widehat{\alpha}_j e^{E[\widehat{Z}(j)]+\frac{1}{2}\sigma^2_{\widehat{Z}(j)}}, \qquad i = 1, ..., n. \tag{29}$$

For this special choice the first order approximation of $Var(V^l)$ is maximized. Therefore this approach will be referred to as *"maximal variance" lower bound approach*. Indeed, from (26) follows:

$$Var(V^l) = \sum_{i=1}^{n}\sum_{j=1}^{n} \widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)})}(e^{r_i r_j \sigma_{\widehat{Z}(i)}\sigma_{\widehat{Z}(j)}}-1)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)})}(r_i r_j \sigma_{\widehat{Z}(i)}\sigma_{\widehat{Z}(j)}) + o(r_i r_j \sigma_{\widehat{Z}(i)}\sigma_{\widehat{Z}(j)})$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \widehat{\alpha}_i\widehat{\alpha}_j e^{E[\widehat{Z}(i)]+E[\widehat{Z}(j)]+\frac{1}{2}(\sigma^2_{\widehat{Z}(i)}+\sigma^2_{\widehat{Z}(j)})}\left(\frac{cov[\widehat{Z}(i),\Lambda]cov[\widehat{Z}(j),\Lambda]}{Var(\Lambda)}\right) + o(r_i r_j \sigma_{\widehat{Z}(i)}\sigma_{\widehat{Z}(j)})$$

$$= \frac{\left(cov\left[\sum_{i=1}^{n}\widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}}\widehat{Z}(i),\Lambda\right]\right)^2}{Var(\Lambda)} + o(r_i r_j \sigma_{\widehat{Z}(i)}\sigma_{\widehat{Z}(j)})$$

$$= \left(r\left[\sum_{i=1}^{n}\widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}}\widehat{Z}(i),\Lambda\right]\right)^2 Var\left[\sum_{i=1}^{n}\widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}}\widehat{Z}(i)\right] + o(r_i r_j \sigma_{\widehat{Z}(i)}\sigma_{\widehat{Z}(j)}).$$

It stands to reason, that the first order approximation of $Var(V^l)$ will be maximized when the correlation coefficient in the last expression is maximized, i.e. when it is taken to be equal to 1. Hence Λ must be chosen as

$$\Lambda = \sum_{i=1}^{n}\widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}}\widehat{Z}(i),$$

which is just the same if to take Λ as

$$\Lambda = \sum_{i=1}^{n}\beta_i\widehat{Z}_i,$$

with coefficients β_i given by (29).

5 Risk measures

5.1 Well-known risk measures

The aim of this section is to describe briefly several well-known risk measures and several important relations which hold between them. One of the most commonly used risk measures in the field of finance and stochastic is the p-quantile risk measure, based on a percentile concept. It is also called *Value-at-Risk* (VaR) at level p in financial literature. For any $p \in (0,1)$, the p-quantile risk measure for a random variable X, denoted by $Q_p[X]$ is a non-decreasing and left-continuous function of p, which is given by

$$Q_p[X] = \inf\{x \in \mathbb{R} | F_X(x) \geq p\}, \tag{30}$$

where by convention $\inf\{\varnothing\} = +\infty$. With the notation of Section 1.2 it holds $Q_p[X] = F_X^{-1}(p)$. Introduce additionally a related risk measure

$$Q_p^+[X] = \inf\{x \in \mathbb{R} | F_X(x) > p\} = \sup\{x \in \mathbb{R} | F_X(x) \leq p\} = F_X^{-1+}(p),$$

where by convention $\sup\{\varnothing\} = -\infty$. $Q_p^+[X]$ is referred to as the *upper p-quantile*, while $Q_p[X]$ is also referred to as *lower p-quantile*. From

$$\{x \in \mathbb{R} | F_X(x) > p\} \subseteq \{x \in \mathbb{R} | F_X(x) \geq p\}$$

follows immediately that $Q_p \leq Q_p^+$. It should be repeated, that only values of p corresponding to horizontal segments of F_X lead to different values of $Q_p[X]$ and $Q_p^+[X]$. Thus if F_X is strictly increasing, both risk measures $Q_p[X]$ and $Q_p^+[X]$ will coincide for all p. In Fig.3 the distribution function of the random variable X with $P[X = -1.5] = 0.1$ and $P[X = 0.5] = 1$ is plotted. Obviously, upper and lower 0.1-quantiles do not coincide.

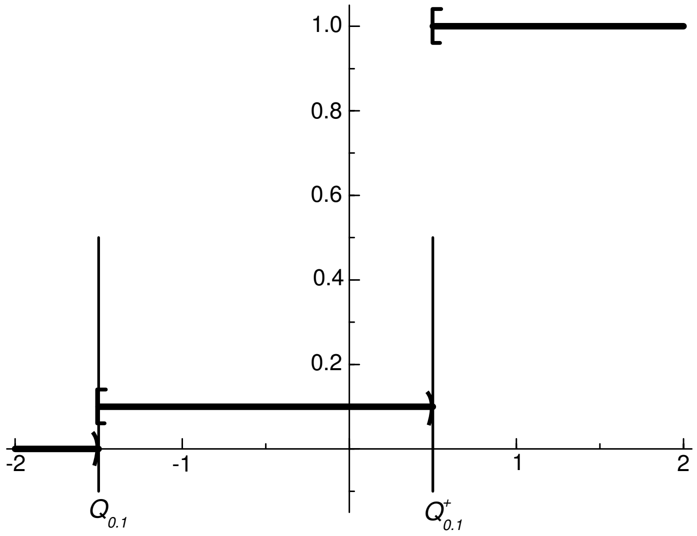

Fig. 3: Difference between the lower and upper quantiles

The concept of Value-at-Risk has become the standard risk measure used to evaluate exposure to risk. Saying it roughly, the VaR is the amount of capital required to ensure, with a high degree of certainty, that the enterprise does not become technically insolvent. The degree of certainty p can be arbitrary chosen. In practice, it can be a high number such as 99.95% for the entire enterprise, or it can be much lower, e.g. 95% for a single unit within the enterprise. In current situation, to estimate $Risk_1$ and $Risk_2$ given by Formulas (1) and (2) respectively we are especially interested in low degrees of certainty such as $p = 0.05$ or $p = 0.01$. It should be noted, that VaR is a risk measure that only concerns about the frequency of default. For instance, doubling the largest loss may not impact the VaR at all. Nevertheless from shareholders and management perspective, the VaR at the company level is a meaningful risk measure since the default event itself is of primary concern, while the size of shortfall is only secondary.

In the frame of this thesis, several risk measures concerning the upper tail of the corresponding distributions are considered. One of them is the *Conditional Tail Expectation* (CTE) at level p. It denoted by $CTE_p[X]$ and defined as

$$CTE_p[X] = E[X|X > Q_p[X]]. \tag{31}$$

CTE measures a mean size of the highest $(1-p)\cdot 100\%$ of the values of X, exceeding the quantile "VaR", but it ignores values below the quantile "VaR". This risk measure reflects not only the frequency of default, but also the expected value of default. It is obvious that CTE will be always larger or equal than VaR measure for the same value of p, since it is the VaR plus the expected loss, i.e.

$$CTE_p[X] = E[X|X > Q_p[X]] = Q_p[X] + E[X - Q_p[X]|X > Q_p[X]].$$

Another example of risk measure is *Tail-Value-at-Risk*($TVaR$) at level p which is defined as

$$TVaR_p[X] = \frac{1}{1-p}\int_p^1 Q_q[X]dq. \tag{32}$$

It can be interpreted as the arithmetic average of the quantiles of X, from p on. Obviously, $TVaR$ is always larger than the corresponding quantile.

The *Expected Shortfall (ESF)* at level p is denoted by $ESF_p[X]$ and is defined as:

$$ESF_p[X] = E[(X - Q_p[X])_+],$$

which is the unconditional expectation of excess losses, taking the value zero when losses are less than the quantile. For example, in insurance terms, it is the expected loss under a stop-loss contract with attachment point Q_p.

The following two important relations hold true between CTE, ESF, $TVaR$ and the quantile risk measure Q_p.

Theorem 5.1: For $p \in (0, 1)$, it holds

1. $$CTE_p[X] = Q_p[X] + \frac{1}{1 - F_X(Q_p[X])}ESF_p[X], \tag{33}$$

2. $$TVaR_p[X] = Q_p[X] + \frac{1}{1-p}ESF_p[X]. \tag{34}$$

Proof.

1.

$$\begin{aligned}
ESF_p[X] &= E[(X - Q_p[X])_+] \\
&= E(X - Q_p[X] | X > Q_p[X]) P(X > Q_p[X]) \\
&= E(X - Q_p[X] | X > Q_p[X])(1 - F_X(Q_p[X])) \\
&= (CTE_p[X] - Q_p[X])(1 - F_X(Q_p[X])). \\
\Longrightarrow CTE_p[X] &= Q_p[X] + \frac{ESF_p[X]}{1 - F_X(Q_p[X])}
\end{aligned}$$

2. Let U be a random variable uniformly distributed on $[0, 1]$. Then it holds

$$\begin{aligned}
ESF_p[X] &= E[(X - Q_p[X])_+] \\
&= E[(F_X^{-1}(U) - Q_p[X])_+] \\
&= \int_0^1 (Q_q[X] - Q_p[X])_+ dq.
\end{aligned}$$

From $Q_q[X] > Q_p[X] \Longleftrightarrow q > p$ immediately follows that

$$\begin{aligned}
ESF_p[X] &= \int_p^1 Q_q[X] dq - Q_p[X](1 - p) \\
&= (TVaR_p[X] - Q_p[X])(1 - p).
\end{aligned}$$

Hence $TVaR$ is given by (34). ⊠

The following elementary result will be used later. If F_X is continuous then

$$CTE_p[X] = TVaR_p[X], \; p \in (0, 1). \tag{35}$$

In the sequel, the results of following lemma will be frequently used.

Lemma 5.1: Consider a random variable X that is lognormally distributed, i.e. $\ln X \sim N(\mu, \sigma^2)$, then quantile and Conditional Tail Expectation of X can be calculated as

$$(1) \;\; Q_p[X] = e^{\mu + \sigma \Phi^{-1}(p)}, \tag{36}$$

$$(2) \;\; CTE_p[X] = e^{\mu + \frac{\sigma^2}{2}} \frac{\Phi(\sigma - \Phi^{-1}(p))}{1 - p}. \tag{37}$$

Proof.

(1) In view of the fact that $Q_p[X] = \inf\{x \in \mathbb{R} | F_X(x) \geq p\} = F_X^{-1}(p)$ the proof of the first assertion coincides with the proof of Lemma 4.1.

(2) The Black-Scholes call-option pricing formula is based on the following expression for the stop-loss premium of lognormal random variable

$$E[(X - d)_+] = e^{\mu + \frac{\sigma^2}{2}} \Phi(d_1) - d\Phi(d_2), \quad d > 0, \tag{38}$$

with $d_1 = \sigma + \frac{\mu - \ln d}{\sigma}$ and $d_2 = d_1 - \sigma$. To be more precise, consider a European call-option with maturity T and exercise price K. The pay-off of call-option at the expiration date is $(S(T) - K)_+$. Then its fair price at time point 0 is given by

$$C_0 = E(e^{-rT}(S(T) - K)_+) = e^{-rT} E((S_0 e^{\left(r - \frac{\nu^2}{2}\right)T + \nu\sqrt{T}\widetilde{z}} - K)_+),$$

where $\widetilde{z}$ is normally (0,1) distributed random variable. Note that

$$S(t) = S_0 \exp\left(\left(r - \frac{\nu^2}{2}\right) t + \nu W_t\right)$$

characterizes the stock price at time t in the risk neutral world and consequently $S(T)$ is lognormally distributed random variable. Applying Formula (38) with $\mu = \ln S_0 + \left(r - \frac{\nu^2}{2}\right) T$, $\sigma = \nu\sqrt{T}$ and $d = K$, one obtains

$$C_0 = S_0 \Phi\left(\frac{\left(r + \frac{\nu^2}{2}\right) T + \ln \frac{S_0}{K}}{\nu\sqrt{T}}\right) - K e^{-rT} \Phi\left(\frac{\left(r - \frac{\nu^2}{2}\right) T + \ln \frac{S_0}{K}}{\nu\sqrt{T}}\right).$$

The last formula is just the famous Black-Scholes formula for pricing an European call-option.

Expected shortfall is defined as stop-loss premium at retention $Q_p[X]$. Thus substituting in Formula (38) $d = Q_p[X]$ with $Q_p[X]$ given by (36) one derives

$$\begin{aligned}
ESF_p[X] &= e^{\mu + \frac{\sigma^2}{2}} \Phi\left(\sigma + \frac{\mu - \mu - \sigma\Phi^{-1}(p)}{\sigma}\right) \\
&- \Phi\left(\sigma + \frac{\mu - \mu - \sigma\Phi^{-1}(p)}{\sigma} - \sigma\right) e^{\mu + \sigma\Phi^{-1}(p)} \\
&= e^{\mu + \frac{\sigma^2}{2}} \Phi(\sigma - \Phi^{-1}(p)) - e^{\mu + \sigma\Phi^{-1}(p)}(1 - p).
\end{aligned}$$

Finally, applying Formula (33) one obtains

$$\begin{aligned}
CTE_p[X] &= e^{\mu + \sigma\Phi^{-1}(p)} + \frac{1}{1 - p}(e^{\mu + \frac{\sigma^2}{2}} \Phi(\sigma - \Phi^{-1}(p)) - e^{\mu + \sigma\Phi^{-1}(p)}(1 - p)) \\
&= e^{\mu + \frac{\sigma^2}{2}} \frac{\Phi(\sigma - \Phi^{-1}(p))}{1 - p}. \quad \boxtimes
\end{aligned}$$

5.2 Distortion risk measures

Using the comonotonic approach one can easily derive the approximate value of any risk measure that is additive for comonotonic risks. Distortion risk measures are examples of such risk measures.

It was Wang (1996) who has first introduced the notion of distortion risk measure in actuarial literature. He defines a class of distortion risk measures by means of the concept of distortion function.

Definition 5.1: Let g: $[0, 1] \to [0, 1]$ be an increasing function with $g(0) = 0$ and $g(1) = 1$. The transform $F^*(x) = g(F_X(x))$ defines a *distorted probability*, where g is called a *distortion function (operator)*.

Definition 5.2: We define *distortion risk measures* using the mean-value under the distorted probability $F^*(x) = g(F_X(x))$

$$\rho_g[X] = E^*[X] = -\int_{-\infty}^{0} [1 - g(\overline{F_X}(x))]dx + \int_{0}^{\infty} g(\overline{F_X}(x))dx. \tag{39}$$

Obviously, a distortion operator g transforms a probability distribution $\overline{F_X}$ into a new probability distribution $g(\overline{F_X})$. Recall, that the expectation of X is given as

$$E[X] = -\int_{-\infty}^{0} [1 - \overline{F_X}(x)]dx + \int_{0}^{\infty} \overline{F_X}(x)dx. \tag{40}$$

It stands to reason, that in the case of identity distortion function $g(x) = x$, the distortion risk measure $\rho_g[X]$ given by (39) turns to be $E[X]$. Thus $E[X]$ is the simplest example of a distortion risk measure. For a general $\rho_g[X]$, Denneberg (1994) and Wang (2000) interpret $\rho_g[X]$ as a "distorted expectation" $E^*[X]$, evaluated with a "distorted probability" in the sense of Choquet integral. Fig. 4 and 5 reflect differences between the usual expectation, given by Formula (40), and "distorted expectation", given by Formula (39).

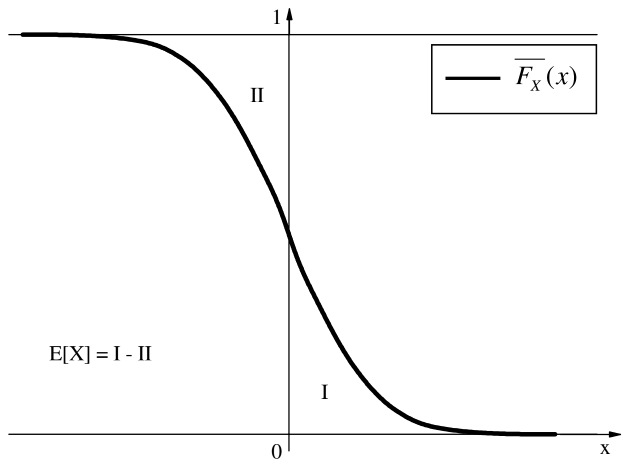

Fig. 4: Expectation

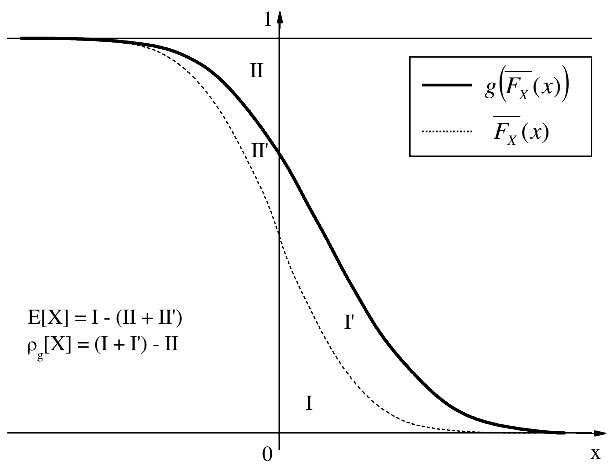

Fig. 5: Distorted expectation

Remark 5.1: A distortion function g is assumed to be independent of the distribution function of the random variable X.

Lemma 5.2: 1. $Q_p[X]$ corresponds to the distortion function

$$g(x) = I_{(x>1-p)}, \quad 0 \le x \le 1.$$

2. $TVaR_p[X]$ corresponds to the distortion function

$$g(x) = \min\left(\frac{x}{1-p}, 1\right), \quad 0 \le x \le 1.$$

3. $CTE_p[X]$ corresponds to the distortion operator

$$g(x) = \min\left(\frac{x}{1-F_X(Q_p[X])}, 1\right), \quad 0 \le x \le 1.$$

Proof.

1. It holds

$$g(\overline{F_X}(x)) = I_{(\overline{F_X}(x)>1-p)} = I_{(1-F_X(x)>1-p)}.$$

From (4) it follows $F_X(x) < p \iff F_X^{-1}(p) > x$, and we find that

$$g(\overline{F_X}(x)) = \begin{cases} 1, & x < F_X^{-1}(p) \\ 0, & x \ge F_X^{-1}(p) \end{cases}$$

Hence, it immediately follows that

$$\rho_g[X] = \int_0^{F_X^{-1}(p)} dx = F_X^{-1}(p) = Q_p[X].$$

2. We have

$$\begin{aligned} g(\overline{F_X}(x)) &= \min\left(\frac{1-F_X(x)}{1-p}, 1\right) \\ &= \begin{cases} 1, & x < F_X^{-1}(p) \\ \dfrac{1-F_X(x)}{1-p}, & x \ge F_X^{-1}(p) \end{cases} \end{aligned}$$

Consequently,

$$\begin{aligned} \rho_g[X] &= \int_0^{F_X^{-1}(p)} dx + \frac{1}{1-p}\int_{F_X^{-1}(p)}^{\infty} \overline{F_X}(x)dx \\ &= \int_0^{F_X^{-1}(p)} dx + \frac{1}{1-p}\int_{F_X^{-1}(p)}^{\infty} (1-F_X(x))dx \\ &= \int_0^{F_X^{-1}(p)} dx + \frac{1}{1-p}\int_{Q_p(X)}^{\infty} (1-F_X(x))dx \\ &= Q_p[X] + \frac{1}{1-p}E[X-Q_p[X]]_+ \\ &= Q_p[X] + \frac{1}{1-p}ESF_p[X] = TVaR_p[X], \end{aligned}$$

where (34) and (8) has been used.

3. From the Formulas (33) and (34) immediately follows that

$$CTE_p[X] = TVaR_{F_X(Q_p[X])}[X].$$

Thus the proof of the assertion 3 immediately follows from the assertion 2. ⊠

Fig. 6 and 7 illustrate the distortion functions for quantile and $TVaR$ respectively. By force of Remark 5.1 Conditional Tail Expectation is not a distortion risk measure as its distortion function depends on the distribution function of X. It was illustrated by Dhaene et al. (2004) that $ESF_p[X]$ cannot be expressed in the form (39) for some distortion function g. Thus it can be concluded that Expected Shortfall is not a distortion risk measure.

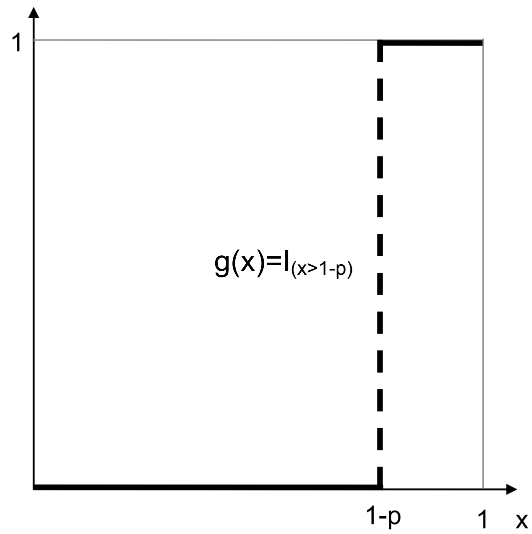

Fig. 6: Distortion function for quantile risk measure

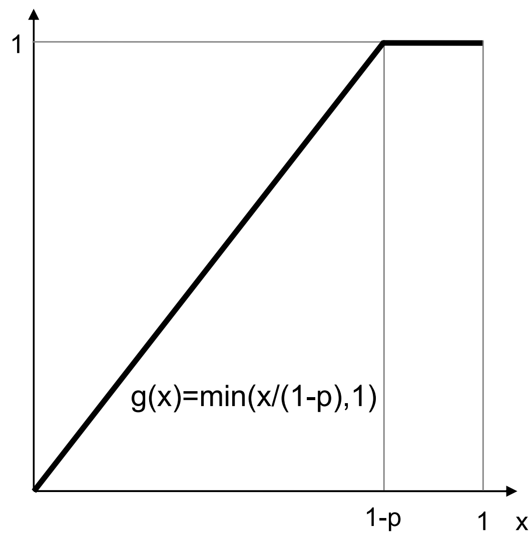

Fig. 7: Distortion function for Tail-Value-at-Risk

Distortion risk measures obey the following properties (the ideas of their proofs can be found in Dhaene et al.(2004)):

1. ***Positive homogeneity***: For any positive constant λ and any distortion function g
$$\rho_g(\lambda x) = \lambda \rho_g(x).$$
If to think of $\rho_g(x)$ as the amount of capital requirement for the risk X, this property means, that the capital requirement is independent of the currency in which risk is measured.

2. ***Additivity for comonotonic risks***: For all random vectors $X = (X_1, X_2, ..., X_n)$ and any distortion function g it holds
$$\rho_g[X_1^c + X_2^c + ... + X_n^c] = \sum_{i=1}^{n} \rho_g[X_i],$$
where $X^c = (X_1^c, X_2^c, ..., X_n^c)$ denotes the comonotonic counterpart of X.
This property means, that the capital requirement for combined risks will be equal to the capital requirements for the risks treated separately.

3. ***Monotonicity***: If $X \leq Y$ for all possible outcomes, then
$$\rho_g[X] \leq \rho_g[Y]$$
for any distortion function g. This implies, that if one risk always has greater loses than another risk, the capital requirement should be greater.

4. ***Translation invariance***: For any positive constant a and any distortion function g
$$\rho_g[X + a] = \rho_g[X] + a.$$
This means, that there is no additional capital requirement for an additional risk for which there is no uncertainty. In particular, making X identically zero, the total capital required for a certain outcome is exactly the value of that outcome.

5.3 Distortion risk measures and comonotonicity

The advantage of the comonotonic dependency structure is that any distortion risk measure of the sum of comonotonic random variables can be easily calculated as the sum of risk measures of the marginals involved. This result, which is especially important for future calculations, is presented in the theorem below. Because the general proof of property 2 for distortion risk measures shall not be a part of this thesis, the proof of the additivity property is given directly for the risk measures under consideration.

Theorem 5.2: Consider a random vector $X = (X_1, X_2, ..., X_n)$ and its comonotonic counterpart $(X_1^c, X_2^c, ..., X_n^c)$. Let
$$V^c = X_1^c + X_2^c + ... + X_n^c.$$
Then for all $p \in (0, 1)$ the following properties hold:
$$Q_p[V^c] = \sum_{i=1}^{n} Q_p[X_i], \tag{41}$$

$$TVaR_p[V^c] = \sum_{i=1}^{n} TVaR_p[X_i]. \tag{42}$$

Proof. In accordance with the Definition 3.5 of the comonotonic counterpart it holds

$$(X_1^c, X_2^c, ..., X_n^c) \stackrel{d}{=} (F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), ..., F_{X_n}^{-1}(U)).$$

Thus

$$V^c \stackrel{d}{=} F_{X_1}^{-1}(U) + F_{X_2}^{-1}(U) + ... + F_{X_n}^{-1}(U) =: g(U).$$

The function $g(\cdot)$ is the sum of non-decreasing and left-continuous functions $F_{X_i}^{-1}(\cdot)$, $i = 1, ..., n$, hence it is non-decreasing and left-continuous function itself. Recall the following relation between the inverse distribution function of the random variable X and non-decreasing and left-continuous function g (see Remark 1.2 (assertion 1)):

$$F_{g(X)}^{-1}(p) = g(F_X^{-1}(p)), \quad p \in (0,1).$$

Since $Q_p[X] = F_X^{-1}(p)$, the last relation is equivalent to

$$Q_p[g(X)] = g(Q_p[X]).$$

Consequently,

$$\begin{aligned} Q_p[V^c] &= Q_p[g(U)] = g(Q_p(U)) \\ &= F_{X_1}^{-1}(Q_p(U)) + F_{X_2}^{-1}(Q_p(U)) + ... + F_{X_n}^{-1}(Q_p(U)). \end{aligned}$$

Obviously, that for a uniform (0,1) random variable U holds $F_U^{-1}(p) = Q_p(U) = p$. Hence

$$\begin{aligned} Q_p[V^c] &= F_{X_1}^{-1}(p) + F_{X_2}^{-1}(p) + ... + F_{X_n}^{-1}(p) \\ &= Q_p[X_1] + Q_p[X_2] + ... + Q_p[X_n]. \end{aligned}$$

This is just the proof of the relation (41).

In order to prove the relation (42) Formula (32) for Tail-Value-at-Risk as well as already proved relation (41) must be simultaneously applied. Namely

$$\begin{aligned} TVaR_p[V^c] &= \frac{1}{1-p}\int_p^1 Q_q[V^c]dq \\ &= \frac{1}{1-p}\int_p^1 \sum_{i=1}^{n} Q_q[X_i]dq \\ &= \sum_{i=1}^{n} \frac{1}{1-p}\int_p^1 Q_q[X_i]dq \\ &= \sum_{i=1}^{n} TVaR_p[X_i]. \ \boxtimes \end{aligned}$$

Remark 5.2: The CTE risk measure is in general not additive for comonotonic risks. However, the additivity property holds for the CTE when the comonotonic random vector $(X_1^c, X_2^c, ..., X_n^c)$ has continuous marginal distributions. In this case V^c can be presented as a strongly monotonically increasing function of (the continuously distributed) random variable U, therefore V^c is also continuously distributed (see Shirayev (1996), p.240). Thus, from (35) follows that $CTE_p[V^c] = TVaR_p[V^c]$ for all $p \in (0,1)$. Consequently, from (42) follows that

$$CTE_p[V^c] = TVaR_p[V^c] = \sum_{i=1}^{n} TVaR_p[X_i] = \sum_{i=1}^{n} CTE_p[X_i].$$

If the marginal distributions are not continuous, CTE is not additive in general.

5.4 Concave distortion risk measures

In the sequel, a special attention will be paid to the so-called class of concave distortion risk measures, which is a subclass of distortion risk measures. It is intuitively clear, that a concave distortion risk measure is defined as a distortion risk measure with a concave distortion function.

Definition 5.3: A distortion function g will said to be *concave* if for each $y \in (0,1]$, there exist real numbers a_y and b_y and a line $l(x) = a_y x + b_y$, such that $l(y) = g(y)$ and $l(x) \geq g(x)$ for all $x \in (0,1]$.

A concave distortion risk measure is necessarily continuous in $(0,1]$. It is also assumed to be continuous at point 0 from the reason of convenience. Fig. 6 and 7 illustrate the distortion functions for quantile and Tail-Value-at-Risk respectively. So it becomes evident that the quantile risk measure is not a concave distortion risk measure while $TVaR$ is a concave distortion risk measure.

The following theorem is especially important to compare the evaluated risk measures for comonotonic lower and upper bounds. It shows that stop-loss order can be characterized in terms of ordered concave distortion risk measures.

Theorem 5.3: For any random pair (X, Y) we have that $X \leq_{sl} Y$ if and only if their respective concave distortion risk measures are ordered

$$X \leq_{sl} Y \Longleftrightarrow \rho_g[X] \leq \rho_g[Y] \tag{43}$$

for all concave distortion functions g.

Proof. See Yaari (1987), Wang and Young (1998) and Dhaene et al. (2000).

If for any random variables X and Y the property (43) holds true, then the risk measure ρ_g is said to *preserve stop-loss order*. As a special case of Theorem 5.3 it can be shown that concave $TVaR$ preserves stop-loss order. It is formulated in the following theorem, which is proved directly.

Theorem 5.4: For any random pair (X, Y) we have that $X \leq_{sl} Y$ if and only if their respective $TVaR$'s are ordered

$$X \leq_{sl} Y \iff TVaR_p[X] \leq TVaR_p[Y]$$

for all $p \in (0,1)$.

Proof (follows the ideas given in Dhaene et all. (2004)).
To prove "$\Longrightarrow$" implication, assume that $X \leq_{sl} Y$ and let $p \in (0,1)$. With the help of relation (9) construct the function

$$f(d) = (1-p)d + E[(X-d)_+] = (1-p)d + \int_d^\infty \overline{F}_X(x)dx.$$

From the monotonicity of the decumulative distribution function $\overline{F}_X(x)$ follows that the function $f(d)$, and consequently $f(d)/(1-p)$, is minimized for $d = Q_p[X]$. Thus

$$\begin{aligned} TVaR_p[X] &= Q_p[X] + \frac{1}{1-p}E[(X-Q_p[X])_+] \\ &= \frac{f(Q_p[X])}{1-p} \leq \frac{f(Q_p[Y])}{1-p} \\ &= Q_p[Y] + \frac{1}{1-p}E[(X-Q_p[Y])_+] \leq TVaR_p[Y], \end{aligned}$$

where the initial condition that $X \leq_{sl} Y$ was used to derive last inequality.
To prove "$\Longleftarrow$" implication, assume that

$$TVaR_p[X] \leq TVaR_p[Y]$$

for all $p \in (0,1)$. It can be derived by analogy with the proof of assertion 2 in Theorem 5.1 that

$$E[(X-d)_+] = E[(F_X^{-1}(U)-d)_+] = \int_0^1 (F_X^{-1}(q)-d)_+ dq.$$

From the fact that $F_X^{-1}(q) > d \Longleftrightarrow q > F_X(d)$ immediately follows

$$\begin{aligned} E[(X-d)_+] &= \int_{F_X(d)}^1 (Q_q[X]-d)dq \\ &= \int_{F_X(d)}^1 Q_q[X]dq - d(1-F_X(d)). \end{aligned} \tag{44}$$

Hence, for d such that $0 < F_X(d) < 1$, we find

$$\begin{aligned} E[(X-d)_+] &= TVaR_{F_X(d)}[X](1-F_X(d)) - d(1-F_X(d)) \\ &= (TVaR_{F_X(d)}[X]-d)(1-F_X(d)) \\ &\leq (TVaR_{F_X(d)}[Y]-d)(1-F_X(d)) \\ &= \int_{F_X(d)}^1 Q_q[Y]dq - d(1-F_X(d)) \\ &= \int_{F_X(d)}^{F_Y(d)} Q_q[Y]dq - d(F_Y(d)-F_X(d)) \\ &\quad + \int_{F_Y(d)}^1 Q_q[Y]dq - d(1-F_Y(d)) \\ &= \int_{F_X(d)}^{F_Y(d)} (Q_q[Y]-d)dq + E[(Y-d)_+]. \end{aligned}$$

Since the equivalence $q \leq F_Y(d) \Longleftrightarrow d \geq Q_p[Y]$ holds true, the following inequality takes place

$$\int_{F_X(d)}^{F_Y(d)} (Q_q[Y] - d)dq \leq 0.$$

This proves that

$$E[(X-d)_+] \leq E[(Y-d)_+]$$

for any retention d such that $0 < F_X(d) < 1$. If $F_X(d) = 1$, find from (44) that

$$E[(X-d)_+] = 0 \leq E[(Y-d)_+].$$

Assume that $F_X(d) = 0$. Using the following elementary result that

$$\lim_{p \to 0} TVaR_p[X] = \lim_{p \to 0} \frac{1}{1-p} \int_p^1 Q_q[X]dq = E[X],$$

one can immediately derive from $TVaR_p[X] \leq TVaR_p[Y]$ that $E[X] \leq E[Y]$. Then from (44) follows that $E[(X-d)_+] \leq E[(Y-d)_+]$ holds for d such that $F_X(d) = 0$. ⊠

Any concave distortion risk measure possesses the **subadditivity property**. This means that the risk measure for a sum of random variables is smaller or equivalent to the sum of the risk measures. It can be formulated as follows

$$\rho_g[X+Y] \leq \rho_g[X] + \rho_g[Y]$$

for any concave distortion function g, and any two random variables X and Y. If to think of $\rho(X)$ as the amount of solvency capital required for the risk X, then subadditivity property has the following financial interpretation: the capital requirement for two risks combined will not be greater, than for the risks treated separately. This is necessary, since otherwise companies would have an advantage to disaggregate into smaller companies.

The proof of the subadditivity property for concave distortion risk measures is a straightforward consequence of the next theorem.

Theorem 5.5: Any risk measure that is additive for comonotonic risks and that preserves stop-loss order is sub-additive.

Proof. We have that a sum of random variables with given marginal distributions is largest in the convex order sense if these random variables are comonotonic (see Theorem 4.1). Convex order implies corresponding stop-loss order, hence

$$X + Y \leq_{sl} X^c + Y^c.$$

If the risk measure ρ preserves stop-loss order and is additive for comonotonic risks, then

$$\rho[X+Y] \leq \rho[X^c + Y^c] = \rho[X] + \rho[Y],$$

which proves the stated result. ⊠

As a special case of Theorem 5.5 immediately follows that $TVaR$ is sub-additive, as it is additive as any distortion risk measure and preserves stop-loss order as any concave distortion risk measure thus

$$TVaR_p[X+Y] \leq TVaR_p[X] + TVaR_p[Y].$$

In its turn CTE, VaR and ESF are not sub-additive (for details see Dhaene et al. (2004)).

Risk measures satisfying sub-additivity, monotonicity, positive homogeneity and translation invariance criteria are deemed to be "Artzner" coherent (see Artzner (1999), Artzner et al. (1999), Panjer et al. (2002) or Wang (2002)). Consequently, any concave distortion measure is coherent.

6 Applications

6.1 Risk measures for sums of dependent lognormal random variables

Consider the general terminal wealth problem as described in Section 1.1. The final wealth, defined by

$$V = \sum_{i=1}^{n} \widehat{\alpha}_i e^{\widehat{Z}(i)} = \sum_{i=1}^{n} X_i$$

with

$$X_i = \widehat{\alpha}_i e^{\widehat{Z}(i)},$$

is a sum of non-independent lognormal random variables. As was already mentioned, it is impossible to determine the distribution function of V analytically. Thus it might be helpful to approximate the distribution function of V by the distribution function of

$$V^c = \sum_{i=1}^{n} F_{X_i}^{-1}(U) \tag{45}$$

or by the distribution function of

$$V^l = \sum_{i=1}^{n} E[X_i|\Lambda],$$

see Theorem 4.1 and Theorem 4.2 for details. Therefore the "distorted expectations" $\rho_g[V]$ of the sum V (such as quantiles and $TVaR$'s) can be approximated by the "distorted expectations" $\rho_g[V^c]$ of upper bound or by the "distorted expectations" $\rho_g[V^l]$ of lower bound.

To be more specific, let us consider the upper bound. In view of Theorem 4.3, the closed-form expression for the comonotonic upper bound is given by

$$V^c = \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\sigma_{\widehat{Z}(i)}\Phi^{-1}(U)}.$$

From the additivity property for comonotonic risks (see Theorem 5.2) it follows

$$\begin{aligned} Q_p[V^c] &= Q_p[\sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\sigma_{\widehat{Z}(i)}\Phi^{-1}(U)}] \\ &= \sum_{i=1}^{n} \widehat{\alpha}_i Q_p[e^{E[\widehat{Z}(i)]+\sigma_{\widehat{Z}(i)}\Phi^{-1}(U)}]. \end{aligned}$$

In the last expression the random variable inside the brackets is lognormally distributed with parameters $E[\widehat{Z}(i)]$ and $\sigma^2_{\widehat{Z}(i)}$. Thus, one can immediately apply Lemma 5.1 to each component in the last sum. This leads to

$$Q_p[V^c] = \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\sigma_{\widehat{Z}(i)}\Phi^{-1}(p)}.$$

By analogy

$$\begin{aligned} CTE_p[V^c] &= CTE_p[\sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\sigma_{\widehat{Z}(i)}\Phi^{-1}(U)}] \\ &= \sum_{i=1}^{n} \widehat{\alpha}_i CTE_p[e^{E[\widehat{Z}(i)]+\sigma_{\widehat{Z}(i)}\Phi^{-1}(U)}] \\ &= \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}} \frac{\Phi(\sigma_{\widehat{Z}(i)} - \Phi^{-1}(p))}{1-p}, \qquad p \in (0,1), \end{aligned} \tag{46}$$

where in deriving Formula (46) the fact was used that the CTE is additive for comonotonic risks with continuous marginal distribution (see Remark 5.2).

In order to define a stochastic lower bound for V, the conditioning random variable Λ, which is linear combination of the Z_j

$$\Lambda = \sum_{j=1}^{n} \beta_j Z_j$$

must be chosen. In particular, coefficients β_j must be chosen in a proper way. As described in Section 4.3, two possibilities for choosing Λ shall be under consideration. The idea of the lower bound approach is to take β_j as

$$\beta_j = \sum_{i=j}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]}. \tag{47}$$

The "maximal variance" lower bound approach is to choose β_j as

$$\beta_j = \sum_{i=j}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}}. \tag{48}$$

After some computations (see Theorem 4.3) it can be found that stochastic lower bound V^l is given by

$$V^l = \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\frac{1}{2}(1-r_i^2)\sigma^2_{\widehat{Z}(i)}+r_i\sigma_{\widehat{Z}(i)}\Phi^{-1}(U)},$$

where the uniformly (0,1) distributed random variable U follows from

$$\Phi^{-1}(U) \equiv \frac{\Lambda - E(\Lambda)}{\sigma_\Lambda}$$

and

$$r_i = \frac{cov(\widehat{Z}(i), \Lambda)}{\sigma_{\widehat{Z}(i)}\sigma_\Lambda}.$$

For the special choices of coefficients β_j given by (47) and (48) the lower bound V^l turns to be comonotonic, since correlation coefficients r_i are positive. This can be seen from the following derivations under the assumption that $t_i = i$:

$$\begin{aligned} E\widehat{Z}(i)E\Lambda &= E\left[\sum_{j=1}^{i} \widehat{Z}_j\right] E\left[\sum_{j=1}^{n} \beta_j \widehat{Z}_j\right] \\ &= i\left(\mu - \frac{\sigma^2}{2}\right)\left(\mu - \frac{\sigma^2}{2}\right)\sum_{j=1}^{n}\beta_j \\ &= i\left(\mu - \frac{\sigma^2}{2}\right)^2 \sum_{j=1}^{n}\beta_j, \end{aligned}$$

$$E(\widehat{Z}(i)\Lambda) = E\left(\sum_{j=1}^{i} \widehat{Z}_j \sum_{k=1}^{n} \beta_k \widehat{Z}_k\right).$$

In the special case that $\widehat{Z}_i$ are independent identically distributed random variables find

$$E(\widehat{Z}(i)\Lambda) = \sum_{j=1}^{i} \beta_j E(\widehat{Z}_j^2) + i\left(\mu - \frac{\sigma^2}{2}\right)^2 \sum_{j=1}^{n}\beta_j - \left(\mu - \frac{\sigma^2}{2}\right)^2 \sum_{j=1}^{i}\beta_j.$$

The second moments of $\widehat{Z}_j$ can be calculated as

$$E(\widehat{Z}_j^2) = Var(\widehat{Z}_j) + (E(\widehat{Z}_j))^2 = \sigma^2 + \left(\mu - \frac{\sigma^2}{2}\right)^2.$$

Consequently

$$\begin{aligned} E(\widehat{Z}(i)\Lambda) &= \left(\sigma^2 + \left(\mu - \frac{\sigma^2}{2}\right)^2\right)\sum_{j=1}^{i}\beta_j + i\left(\mu - \frac{\sigma^2}{2}\right)^2 \sum_{j=1}^{n}\beta_j - \left(\mu - \frac{\sigma^2}{2}\right)^2 \sum_{j=1}^{i}\beta_j \\ &= \sigma^2 \sum_{j=1}^{i}\beta_j + i\left(\mu - \frac{\sigma^2}{2}\right)^2 \sum_{j=1}^{n}\beta_j, \end{aligned}$$

$$\begin{aligned} cov(\widehat{Z}(i),\Lambda) &= E(\widehat{Z}(i)\Lambda) - E\widehat{Z}(i)E\Lambda \\ &= \sigma^2 \sum_{j=1}^{i}\beta_j + i\left(\mu - \frac{\sigma^2}{2}\right)^2 \sum_{j=1}^{n}\beta_j - i\left(\mu - \frac{\sigma^2}{2}\right)^2 \sum_{j=1}^{n}\beta_j \\ &= \sigma^2 \sum_{j=1}^{i}\beta_j. \end{aligned}$$

Thus

$$r_i = \frac{cov(\widehat{Z}(i),\Lambda)}{\sigma_{\widehat{Z}(i)}\sigma_\Lambda} = \frac{\sigma^2 \sum_{j=1}^{i}\beta_j}{\sigma\sqrt{i}\sigma\sqrt{\sum_{j=1}^{n}\beta_j^2}} = \frac{\sum_{j=1}^{i}\beta_j}{\sqrt{i\sum_{j=1}^{n}\beta_j^2}}.$$

The comonotonicity of V^l implies that quantiles and Conditional Tail Expectation (which is the same as Tail-Value-at-Risk in current situation) can be computed by summing the risk measures for the marginals involved. In view of Lemma 5.1 the following formulas for quantiles and Conditional Tail Expectation related to V^l can be easily derived by analogy with Conditional Tail Expectation and quantile related to V^c:

$$Q_p[V^l] = \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\frac{1}{2}(1-r_i^2)\sigma^2_{\widehat{Z}(i)}+r_i\sigma_{\widehat{Z}(i)}\Phi^{-1}(p)},$$

$$CTE_p[V^l] = \sum_{i=1}^{n} \widehat{\alpha}_i e^{E[\widehat{Z}(i)]+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}} \frac{1-\Phi(r_i\sigma_{\widehat{Z}(i)} - \Phi^{-1}(p))}{1-p}, \quad p \in (0,1).$$

It should be noted, that in the case of a concave distortion function g, $\rho_g[V^c]$ is an upper bound while $\rho_g[V^l]$ is lower bound for $\rho_g[V]$, i.e.

$$\rho_g[V^l] \le \rho_g[V] \le \rho_g[V^c]$$

(for details see Theorem 5.3). In general case Conditional Tail Expectation is not a concave distortion risk measure. If marginal distributions of the sum V are continuous CTE turns to be concave distortion risk measure since it coincides with Tail-Value-at-Risk (see Remark 5.2). This implies

$$CTE_p[V^l] \le CTE_p[V] \le CTE_p[V^c].$$

The quantiles of V^l, V and V^c are not necessarily ordered in the same way, as quantile is not a concave distortion risk measure.

6.2 Auxiliary calculations

Recall, that the main object of this thesis is the estimation of risk measures $Risk_1$ and $Risk_2$ given by Formulas (1) and (2) respectively. Obviously, these definitions are slightly different from the risk measures presented in Section 5.1 by Formulas (30), (31). For example, Conditional Left Tail Expectation given by (2) is indifferent towards the values which are above the border $Q_p(X)$, while the opposite situation is observed in (31), where the values below $Q_p(X)$ are ignored. Hence the aim of this section is to establish interrelations between Formulas (1) and (30), (2) and (31) respectively.

It holds

$$\begin{aligned} Risk_2[X] &= CLTE_p(X-b) := -E(X-b|X \le Q_p(X)) \\ &= -E(X|X \le Q_p(X)) + b. \end{aligned}$$

$$\begin{aligned} E[X] &= E(X\ I_{(X \le Q_p(X))} + X\ I_{(X > Q_p(X))}) \\ &= E(X|X \le Q_p(X))P(X \le Q_p(X)) + E(X|X > Q_p(X))P(X > Q_p(X)) \\ &= pE(X|X \le Q_p(X)) + (1-p)E(X|X > Q_p(X)). \end{aligned}$$

Thus

$$E(X|X \le Q_p(X)) = \frac{1}{p}(E[X] - (1-p)E(X|X > Q_p(X))).$$

Consequently

$$CLTE_p(X-b) = \frac{1-p}{p}E(X|X > Q_p(X)) - \frac{E[X]}{p} + b. \tag{49}$$

Formula (49) is the idea of calculating $Risk_2$, given (31).

By analogy the interrelation between (1) and (30) can be established. In accordance with Formula (1)

$$Risk_1[X] = VaR_p(X - b) := -Q_p(X - b).$$

Applying Remark 1.2 (assertion 1) with non-decreasing and left-continuous function $g(x) = x - b$, one obtains:

$$F^{-1}_{g(X)}(p) = Q_p(X - b) = F^{-1}_X(p) - b = Q_p(X) - b.$$

Consequently,

$$Risk_1[X] = VaR_p(X - b) = b - Q_p(X). \tag{50}$$

The last formula gives the idea of calculating (1) on the basis of (30).

7 Moment matching approximations

The aim of this section is to describe briefly two well-known moment matching approximations, which are widely used by practitioners: the reciprocal Gamma and the lognormal moment matching approximations. Moment matching methods approximate the unknown distribution function by a given one such that the first two moments coincide. Both moment matching approximations are originally considered in the case of continuous setting (see Yor (2001)). It would be also interesting to check their accuracy in the case of discrete settings, which will be done later in Section 8 "Numerical illustration".

7.1 The reciprocal Gamma approximation

A Gamma distributed random variable X has the probability density function

$$g(x;\alpha,\beta) := \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, \quad \forall\ x > 0,$$

with parameters $\alpha > 0$ and $\beta > 0$. $\Gamma(\alpha)$ is the Gamma function which satisfies the following equality

$$(\alpha-1)\Gamma(\alpha-1) = \Gamma(\alpha).$$

The cumulative distribution function of X is denoted by $G(x;\alpha,\beta)$ and is available in most statistical software packages. The following easy-to-derive identity is used later to prove some of the results

$$g(x;\alpha,\beta) = \frac{x}{\beta(\alpha-1)} g(x;\alpha-1,\beta), \quad \forall\ \alpha > 1. \tag{51}$$

Definition 7.1: A random variable is reciprocal (or inverse) Gamma distributed if its inverse is Gamma distributed.

Thus consider the random variable $Y = \frac{1}{X}$ which is obviously reciprocal Gamma distributed, under the initial assumption that X is Gamma distributed. The probability density function and cumulative density function of this random variable are given as:

$$g_R(y;\alpha,\beta) = \frac{g(\frac{1}{y};\alpha,\beta)}{y^2}, \quad \forall\ y > 0$$

$$\begin{aligned} G_R(y;\alpha,\beta) &= P(Y \le y) = P\left(\frac{1}{Y} \ge \frac{1}{y}\right) \\ &= P\left(\frac{1}{Y} > \frac{1}{y}\right) = 1 - G\left(\frac{1}{y};\alpha,\beta\right),\ \forall\ y > 0. \end{aligned}$$

The first two moments of the reciprocal Gamma distribution are:

$$M_1 := E[Y] = \frac{1}{\beta(\alpha-1)}, \quad M_2 := E[Y^2] = \frac{1}{\beta^2(\alpha-1)(\alpha-2)}$$

(see Milevsky and Posner (1998)).

Thus the parameters α and β can be easily expressed in terms of the first two moments:

$$\alpha = \frac{2M_2 - M_1^2}{M_2 - M_1^2}, \quad \beta = \frac{M_2 - M_1^2}{M_2 M_1}. \tag{52}$$

Lemma 7.1: The quantiles and Conditional Tail Expectations of a reciprocal Gamma distributed random variable Y are:

$$1. \quad Q_p[Y] = \frac{1}{G^{-1}(1-p;\alpha,\beta)},\ p \in (0,1), \tag{53}$$

$$2. \quad CTE_p[Y] = \frac{G(G^{-1}(1-p;\alpha,\beta);\alpha-1,\beta)}{\beta(1-p)(\alpha-1)},\ p \in (0,1). \tag{54}$$

Proof.
1. To prove the first statement the definition of the quantile and the definition of the cumulative distribution function for reciprocal Gamma distribution must be consistently applied, i.e.

$$\begin{aligned}
Q_p[Y] &= \inf\{y \in \mathbb{R} | F_Y(y) \geq p\} \\
&= \inf\{y \in \mathbb{R} | G_R(y;\alpha,\beta) \geq p\} \\
&= \inf\left\{y \in \mathbb{R} \middle| 1 - G\left(\frac{1}{y};\alpha,\beta\right) \geq p\right\} \\
&= \inf\left\{y \in \mathbb{R} \middle| G\left(\frac{1}{y};\alpha,\beta\right) \leq 1-p\right\} \\
&= \inf\left\{y \in \mathbb{R} \middle| \frac{1}{y} \leq G^{-1}(1-p;\alpha,\beta)\right\} \\
&= \inf\left\{y \in \mathbb{R} \middle| y \geq \frac{1}{G^{-1}(1-p;\alpha,\beta)}\right\} \\
&= \frac{1}{G^{-1}(1-p;\alpha,\beta)}.
\end{aligned}$$

2. In order to derive Formula (54) for the Conditional Tail Expectation the stop-loss premium will be calculated as

$$\begin{aligned}
E[(Y-d)_+] &= \int_d^\infty (y-d) g_R(y;\alpha,\beta)dy \\
&= \int_0^{\frac{1}{d}} \frac{1}{x^2}\left(\frac{1}{x} - d\right) g_R\left(\frac{1}{x};\alpha,\beta\right) dx \\
&= \int_0^{\frac{1}{d}} \left(\frac{1}{x} - d\right) g(x;\alpha,\beta)dx \\
&= \int_0^{\frac{1}{d}} \frac{g(x;\alpha,\beta)}{x} dx - d\int_0^{\frac{1}{d}} g(x;\alpha,\beta)dx.
\end{aligned}$$

Applying Formula (51) the expression for stop-loss premium can be derived, i.e.

$$\begin{aligned}
E[(Y-d)_+] &= \frac{1}{\beta(\alpha-1)} \int_0^{\frac{1}{d}} g(x;\alpha-1,\beta)dx - d\int_0^{\frac{1}{d}} g(x;\alpha,\beta)dx \\
&= \frac{1}{\beta(\alpha-1)} G\left(\frac{1}{d};\alpha-1,\beta\right) - dG\left(\frac{1}{d};\alpha,\beta\right).
\end{aligned}$$

Thus, the Expected Shortfall is equal to

$$\begin{aligned} ESF_p[Y] &= E[(Y - Q_p[Y])_+] = \frac{1}{\beta(\alpha-1)} G(G^{-1}(1-p;\alpha,\beta);\alpha-1,\beta) \\ & \quad - \frac{1}{G^{-1}(1-p;\alpha,\beta)} G(G^{-1}(1-p;\alpha,\beta);\alpha,\beta) \\ &= \frac{1}{\beta(\alpha-1)} G(G^{-1}(1-p;\alpha,\beta);\alpha-1,\beta) - \frac{1-p}{G^{-1}(1-p;\alpha,\beta)}. \end{aligned}$$

Substitute $Q_p[Y]$ and $ESF_p[Y]$ to the general Formula (33) as was already done in Lemma 5.1. Thus

$$\begin{aligned} CTE_p[Y] &= \frac{1}{G^{-1}(1-p;\alpha,\beta)} \\ & \quad + \frac{1}{1-p} \left(\frac{1}{\beta(\alpha-1)} G(G^{-1}(1-p;\alpha,\beta);\alpha-1,\beta) - \frac{1-p}{G^{-1}(1-p;\alpha,\beta)} \right) \\ &= \frac{G(G^{-1}(1-p;\alpha,\beta);\alpha-1,\beta)}{\beta(\alpha-1)(1-p)}. \boxtimes \end{aligned}$$

To approximate the random variable defined by (3) use equations (52) to locate the appropriate density parameters. Consequently, the quantilies and the Conditional Tail Expectations can be easily calculated by (53) and (54) respectively. To calculate $Risk_1$ and $Risk_2$ given by Formulas (1) and (2) the auxiliary Formulas (50) and (49) must be applied.

A theoretical justification for using reciprocal Gamma approximation lies in the fact that the sum of correlated lognormal random variables converges to the reciprocal Gamma distribution under suitable conditions. This important theoretical result is contained in the following theorem.

Theorem 7.1: Define the integral

$$I_{[0,T]} = \int_0^T \exp\left\{ \left(\mu - \frac{1}{2}\sigma^2\right) s + \sigma B_s \right\} ds.$$

If $\mu - \frac{1}{2}\sigma^2 < 0$, then

$$I_\infty := \lim_{T\to\infty} I_{[0,T]}$$

is reciprocal Gamma distributed with parameters $\alpha = 1 - \frac{2\mu}{\sigma^2}$ and $\beta = \frac{\sigma^2}{2}$, which implies that

$$\lim_{T\to\infty} P[I_{[0,T]} \geq \omega] = \lim_{T\to\infty} P\left[\frac{1}{I_{[0,T]}} \leq \frac{1}{\omega}\right] = G\left(\frac{1}{\omega};\alpha,\beta\right).$$

The proof of this elegant fact is given in Milevsky (1997) or Milevsky and Posner (1998); it was obtained by using scale functions and martingale techniques from stochastic calculus. It should be also remarked that another formulation of this basic result in terms of Bessel processes is contained in Yor (2001) or Dufresne (2004a).

7.2 The lognormal approximation

A lognormal distributed random variable X has the probability density function

$$f_X(x;\mu,\sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}},\ x>0,$$

with parameters $\mu \in \mathbb{R}$, $\sigma > 0$.

The first two moments of X are:

$$M_1 := E[X] = e^{\mu+\frac{1}{2}\sigma^2}, \qquad M_2 := E[X^2] = e^{2\mu+2\sigma^2} \tag{55}$$

Inverting the parameters α and β in terms of the first two moments leads to:

$$\mu = \ln\left(\frac{M_1^2}{\sqrt{M_2}}\right), \tag{56}$$

$$\sigma^2 = \ln\left(\frac{M_2}{M_1^2}\right). \tag{57}$$

It was already proved in Lemma 5.1 that quantiles and Conditional Tail Expectations of X, which is lognormally distributed, have the following form:

$$Q_p[X] = e^{\mu+\sigma\Phi^{-1}(p)},\ p\in(0,1) \tag{58}$$

$$CTE_p[X] = e^{\mu+\frac{\sigma^2}{2}}\frac{\Phi(\sigma-\Phi^{-1}(p))}{1-p},\ p\in(0,1) \tag{59}$$

Thus by analogy with reciprocal Gamma approximation it is possible to approximate the random variable V by lognormal distribution with moments (55). Then substituting coefficients (56) and (57) into Formulas (58) and (59) one can obtain the values of quantiles and Conditional Tail Expectations. The main object of this thesis is to calculate $Risk_1$ and $Risk_2$ given by Formulas (1) and (2) respectively. The values of these risk measures can be easily obtained applying auxiliary Formulas (50) and (49). The theoretical background for lognormal approximation was given by Dufresne (2004b) who obtains a lognormal distribution for sums of dependent lognormal random variables as volatility tends to zero.

8 Numerical illustration

In this section the accuracy and efficiency of lognormal (LN), reciprocal Gamma (RG), comonotonic upper bound (UB), comonotonic lower bound (LB) and "maximal variance" lower bound (MVLB) approximations will be examined. In order to judge their accurateness, the results of this approximations for quantiles and Conditional Left Tail Expectations given by Formulas (1) and (2) respectively will be compared with results obtained by Monte Carlo simulation.

Consider random variable V which is defined in current situation as the random final value of accumulating *unit* saving amounts due at time $0, 1, ..., n-1$, i.e.

$$V = \sum_{i=0}^{n-1} e^{Z_i + Z_{i+1} + ... + Z_{n-1}} = \sum_{k=1}^{n} e^{\widehat{Z}_1 + \widehat{Z}_2 + ... + \widehat{Z}_k} = \sum_{k=1}^{n} e^{\widehat{Z}(k)}$$

(see Section 1.1 for details, as V is a special case of the sum of the general form (3)). The multivariate distribution function of the random vector $(\widehat{Z}_1, \widehat{Z}_2, ..., \widehat{Z}_n)$ is completely specified. In particular, $\widehat{Z}_i$ are independent identically distributed random variables according to $N\left(\mu - \frac{\sigma^2}{2}, \sigma^2\right)$. This implies that

$$E[\widehat{Z}(i)] = i\left(\mu - \frac{\sigma^2}{2}\right),$$
$$Var[\widehat{Z}(i)] = i\sigma^2.$$

In order to compute the lower bound approximations the conditioning random variable Λ is defined as before

$$\Lambda = \sum_{i=1}^{n} \beta_i \widehat{Z}_i.$$

Thus one can find that

$$E[\Lambda] = \left(\mu - \frac{\sigma^2}{2}\right) \sum_{k=1}^{n} \beta_k,$$
$$Var[\Lambda] = \sigma^2 \sum_{k=1}^{n} \beta_k.$$

The correlation coefficients r_i can be calculated as

$$r_i = \frac{cov(\widehat{Z}(i), \Lambda)}{\sigma_{\widehat{Z}(i)} \sigma_\Lambda} = \frac{\sum_{k=1}^{i} \beta_k}{\sqrt{i \sum_{k=1}^{n} \beta_k^2}}$$

(the derivation of the last formula can be found in Section 6.1).

The *comonotonic lower bound* approximations for $Risk_1$ and $Risk_2$ will be computed for the following choice of coefficients β_j:

$$\beta_j = \sum_{i=j}^{n} e^{E[\widehat{Z}(i)]} = \sum_{i=j}^{n} e^{i\left(\mu - \frac{\sigma^2}{2}\right)}.$$

The *comonotonic "maximal variance" lower bound* approximations for "distorted expectations" will be obtained for the following choice of parameters β_j:

$$\beta_j = \sum_{i=j}^{n} e^{E[\widehat{Z}(i)]+\frac{1}{2}\sigma^2_{\widehat{Z}(i)}} = \sum_{i=j}^{n} e^{i\mu}.$$

Under assumption that accumulating is performed with deterministic interest rate r, the quantity

$$b = \sum_{i=0}^{n-1} e^{r(n-i)}$$

will serve as a benchmark, which is essential to calculate $Risk_1$ and $Risk_2$ given by Formulas (1) and (2) respectively. This special choice of a benchmark is argued by the fact that, normally, investor expects to get a profit which is at least the sum of the accumulated invested amounts (otherwise he/she will have no incentive to make investments).

The numerous numerical results will be organized in the following way. Firstly, the approximated and simulated quantiles will be computed for the following choice of the parameters:

$$\begin{aligned} n &= 40, \\ p &= 0.05, \\ \mu &= 0.05, \\ \sigma &= 0.15, \\ r &= 0.04. \end{aligned} \tag{60}$$

Afterwards, the sensitivity of approximations with respect to each of the given above parameters will be tested. All the numerical results will be organized in the form of tables. Each cell of these tables corresponding to some analytical approach will involve approximated values for the "distorted expectation" as well as their relative deviations from Monte Carlo simulation. These deviations will be computed as:

$$\frac{Risk_i[V^{appr}] - Risk_i[V^{MC}]}{|Risk_i[V^{MC}]|} \times 100\%, \quad i = 1, 2.$$

In the last formulas V^{appr} is the value obtained by one of the approximations and V^{MC} denotes the simulation result. The last line in each table will display the results of Monte Carlo simulation, which are based on generating 500000 random paths. The approximated/simulated values obtained for initial set of parameters (60) will be distinguished by italics in each of the tables, while the best approximations will be written in bold figures.

The quality of the inverse Gamma and lognormal approximations for the special choice of parameters (60) is illustrated in Fig. 8. One can see that the fit of both approximations is not very good. This visual observation is confirmed by numerical results. See, for example, corresponding cells in Tables 1 and 2 or in any Table given below.

Tables 1 and 2 display approximated and simulated values of $Risk_1[V]$ and $Risk_2[V]$ respectively for different volatility levels. The lower bound approach (LB) is appeared to be the best among other approximations. Even though the performances of lower bound approach become worse for higher volatilities, the relative deviations for $\sigma = 0.35$ are still quite small (less then 1% for both approximated quantiles and $CLTE$'s).

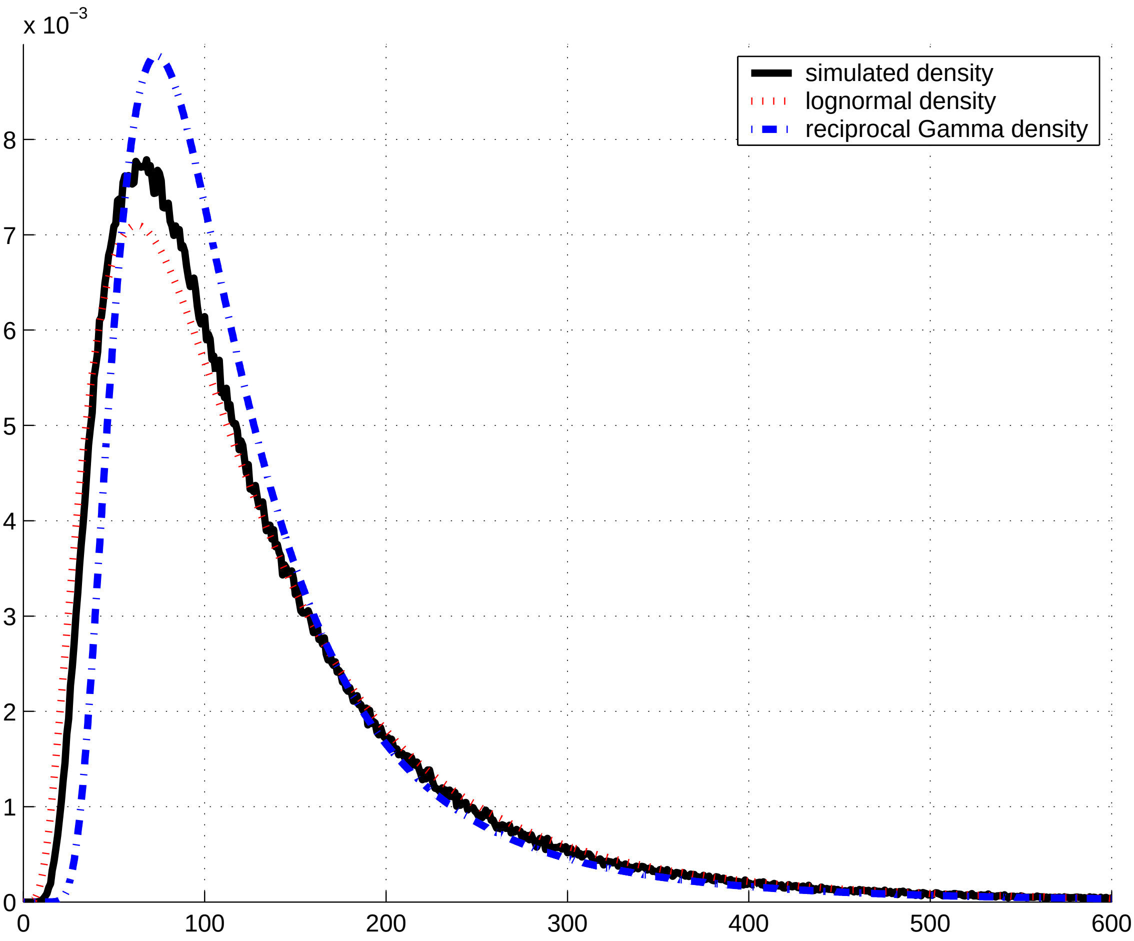

Fig. 8: Simulated and analytical densities for the final value of the cash flow ($n = 40$)

Tab. 1: Approximations for $Risk_1$ for different volatility levels

	Method	$\sigma = 0.05$	$\sigma = 0.15$	$\sigma = 0.25$	$\sigma = 0.35$
n=40	UB	16.494 (+30.75)	*69.890 (+9.70)*	89.902 (+5.62)	96.445 (+3.13)
	LB	**12.571 (-0.35)**	***63.433 (-0.44)***	**84.539 (-0.68)**	**92.843 (-0.72)**
	MVLB	12.568 (-0.37)	*63.287 (-0.67)*	83.892 (-1.44)	91.524 (-2.13)
	RG	11.047 (-12.43)	*53.715 (-15.70)*	68.362 (-19.68)	72.446 (-22.53)
	LN	13.277 (+5.25)	*68.675 (+7.78)*	92.489 (+8.66)	99.435 (+6.33)
	MC	12.616	*63.716*	85.116	93.516

Tab. 2: Approximations for $Risk_2$ for different volatility levels

	Method	$\sigma = 0.05$	$\sigma = 0.15$	$\sigma = 0.25$	$\sigma = 0.35$
n=40	UB	24.333 (+21.89)	*76.592 (+8.36)*	92.885 (+4.66)	97.693 (+2.49)
	LB	**19.925 (-0.19)**	***70.354 (-0.47)***	**88.095 (-0.73)**	**94.588 (-0.76)**
	MVLB	19.921 (-0.21)	*70.177 (-0.72)*	87.433 (-1.48)	93.351 (-2.06)
	RG	17.787 (-10.90)	*60.523 (-14.38)*	73.778 (-16.87)	77.354 (-18.84)
	LN	20.993 (+5.16)	*76.127 (+7.70)*	95.379 (+7.47)	100.044 (+4.95)
	MC	19.963	*70.686*	88.746	95.319

Tables 3 and 4 compare the performances of different approximations for $Risk_1$ and $Risk_2$ in the case of changing interest rates, i.e. in the case of changing benchmarks. Again the lower bound approach approximates both quantiles and Conditional Left Tail Expectations well.

Tab. 3: Approximations for $Risk_1$ for different interest rates

	Method	$r = 0.01$	$r = 0.02$	$r = 0.03$	$r = 0.04$	$r = 0.05$
n=40	UB	18.503 (+48.87)	30.966 (+24.90)	47.577 (+15.19)	*69.890 (+9.70)*	100.076 (+6.58)
	LB	**12.046 (-3.08)**	**24.509 (-1.14)**	**41.120 (-0.44)**	***63.433 (-0.44)***	**93.620 (-0.30)**
	MVLB	11.900 (-4.25)	24.363 (-1.73)	40.975 (-0.79)	*63.287 (-0.67)*	93.474 (-0.46)
	RG	2.329 (-81.27)	14.792 (-40.34)	31.403 (-23.97)	*53.715 (-15.70)*	83.902 (-10.65)
	LN	17.289 (+39.10)	29.752 (+20.01)	46.363 (+12.25)	*68.675 (+7.78)*	98.862 (+5.28)
	MC	12.429	24.791	41.303	*63.716*	93.902

Tab. 4: Approximations for $Risk_2$ for different interest rates

	Method	$r = 0.01$	$r = 0.02$	$r = 0.03$	$r = 0.04$	$r = 0.05$
n=40	UB	25.205 (+30.12)	37.668 (+18.54)	54.279 (+12.44)	*76.592 (+8.36)*	106.779 (+5.86)
	LB	**18.967 (-2.09)**	**31.430 (-1.09)**	**48.041 (-0.48)**	***70.354 (-0.47)***	**100.540 (-0.33)**
	MVLB	18.790 (-3.00)	31.253 (-1.65)	47.864 (-0.85)	*70.177 (-0.72)*	100.364 (-0.50)
	RG	9.136 (-52.83)	21.599 (-32.03)	38.210 (-20.85)	*60.523 (-14.38)*	90.710 (-10.07)
	LN	24.740 (+27.72)	37.203 (+17.08)	53.815 (+11.48)	*76.127 (+7.70)*	106.314 (+5.39)
	MC	19.371	31.777	48.274	*70.686*	100.872

In Tables 5 and 6 the sensitivity of different approximations for $Risk_1$ and $Risk_2$ with respect to yearly horizon n is examined. The numerical results show that the higher the time horizon n, the worse the performances of lognormal and reciprocal Gamma approximations. This numerical observation can be confirmed by Fig. 9, plotting the simulated as well as lognormal and reciprocal densities for time horizon $n = 10$. One can see that the densities for the two moment-matching approximations almost coincide with the simulated one. On the contrary, the increasing of the number of years n in general positively impacts the results of lower bound approach, which is again the best among other methods.

Tab. 5: Approximations for $Risk_1$ for different time horizons

Method	$n = 10$	$n = 20$	$n = 40$	$n = 100$
UB	5.422 (+11.95)	17.201 (+11.78)	*69.890 (+9.70)*	1207.522 (+4.46)
LB	**4.793 (-1.03)**	**15.302 (-0.56)**	***63.433 (-0.44)***	**1150.912 (-0.43)**
MVLB	4.791 (-1.07)	15.285 (-0.67)	*63.287 (-0.67)*	1147.639 (-0.72)
RG	4.555 (-5.94)	14.097 (-8.40)	*53.715 (-15.70)*	641.959 (-44.46)
LN	4.968 (+2.59)	16.230 (+5.47)	*68.675 (+7.78)*	1215.387 (+5.14)
MC	4.843	15.389	*63.716*	1155.931

Tab. 6: Approximations for $Risk_2$ for different time horizons

Method	$n = 10$	$n = 20$	$n = 40$	$n = 100$
UB	6.289 (+12.04)	19.461 (+10.89)	*76.592 (+8.36)*	1260.853 (+3.46)
LB	**5.611 (-0.04)**	**17.501 (-0.28)**	***70.354 (-0.47)***	**1213.853 (-0.39)**
MVLB	5.608 (-0.09)	17.478 (-0.41)	*70.177 (-0.72)*	1210.748 (-0.65)
RG	5.305 (-5.50)	16.125 (-8.12)	*60.523 (-14.38)*	764.058 (-37.30)
LN	5.853 (+4.28)	18.667 (+6.36)	*76.127 (+7.70)*	1270.302 (+4.24)
MC	5.613	17.550	*70.686*	1218.633

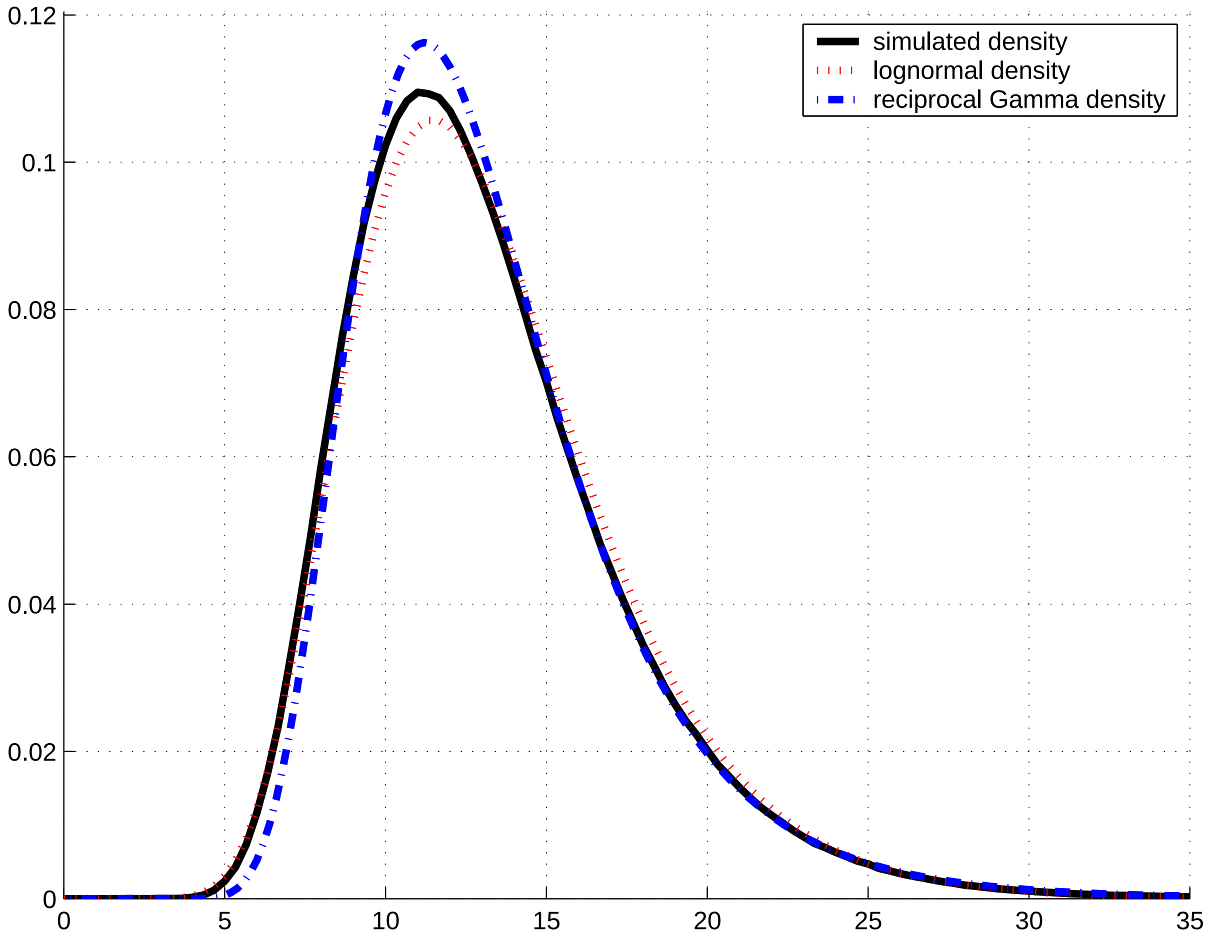

Fig. 9: Simulated and analytical densities for the final value of the cash flow ($n = 10$)

Table 7 and 8 present the approximated/simulated values of $Risk_1$ and $Risk_2$ respectively for different drifts. Observing results presented in these tables one can conclude that deviations of lower bound approach becomes rougher for higher drifts μ. For example, in the case of $\mu = 0.10$ the lower bound approximation for $CLTE$ is extremely poor.

Tab. 7: Approximations for $Risk_1$ for different drifts

	Method	$\mu = 0.05$	$\mu = 0.075$	$\mu = 0.10$
n=40	UB	*69.890 (+9.70)*	45.360 (+28.81)	-6.804 (+71.63)
	LB	***63.433 (-0.44)***	**34.849 (-1.04)**	**-24.689 (-2.94)**
	MVLB	*63.287 (-0.67)*	34.652 (-1.60)	-24.962 (-4.08)
	RG	*53.715 (-15.70)*	10.365 (-70.57)	-85.322 (-255.75)
	LN	*68.675 (+7.78)*	42.980 (+22.05)	-11.937 (+50.23)
	MC	*63.716*	35.215	-23.984

Tab. 8: Approximations for $Risk_2$ for different drifts

	Method	$\mu = 0.05$	$\mu = 0.075$	$\mu = 0.10$
n=40	UB	*76.592 (+8.36)*	58.297 (+19.43)	19.763 (+401.98)
	LB	***70.354 (-0.47)***	**48.336 (-0.97)**	**3.156 (-19.84)**
	MVLB	*70.177 (-0.72)*	48.103 (-1.45)	2.842 (-27.81)
	RG	*60.523 (-14.38)*	23.751 (-51.34)	-57.326 (-1556.08)
	LN	*76.127 (+7.70)*	57.096 (+16.97)	16.621 (+322.17)
	MC	*70.686*	48.811	3.937

Tables 9 and 10 compare different approximations for some selected quantiles and $CLTE$'s respectively. It is an interesting fact, that for high levels of p such as $p = 0.95$ or $p = 0.99$ the maximal variance lower bound approach might be more accurate than lower bound approach. Nevertheless, the lower bound approach outperforms all the other methods when approximating both 0.01-quantiles, 0.01-$CLTE$'s and 0.05-quantiles, 0.05-$CLTE$'s, which are especially important in current situation.

Tab. 9: Approximations for some selected $Risk_1$

	Method	$p = 0.01$	$p = 0.05$	$p = 0.5$	$p = 0.95$	$p = 0.99$
n=40	UB	80.892 (+7.40)	*69.890 (+9.70)*	3.168 (+367.48)	-239.658 (-9.99)	-483.081 (-11.98)
	LB	**74.796 (-0.69)**	***63.433 (-0.44)***	**-1.180 (+0.34)**	**-219.421 (-0.71)**	-428.575 (+0.65)
	MVLB	74.599 (-0.95)	*63.287 (-0.67)*	-1.086 (+8.28)	-219.524 (-0.75)	**-429.794 (+0.37)**
	RG	64.889 (-13.84)	*53.715 (-15.70)*	-2.927 (-147.21)	-199.467 (+8.45)	-420.721 (+2.47)
	LN	80.919 (+7.44)	*68.675 (+7.78)*	-1.454 (-22.80)	-224.603 (-3.08)	-424.863 (+1.51)
	MC	75.315	*63.716*	-1.184	-217.884	-431.384

Tab. 10: Approximations for some selected $Risk_2$

	Method	$p = 0.01$	$p = 0.05$	$p = 0.5$	$p = 0.95$	$p = 0.99$
n=40	UB	84.359 (+6.67)	*76.592 (+8.36)*	40.890 (+16.59)	-10.807 (+15.64)	-23.469 (+3.42)
	LB	**78.506 (-0.73)**	***70.354 (-0.47)***	**35.019 (-0.15)**	-13.087 (-2.16)	-24.379 (-0.33)
	MVLB	78.296 (-0.99)	*70.177 (-0.72)*	34.990 (-0.24)	**-13.039 (-1.79)**	**-24.350 (-0.21)**
	RG	68.854 (-12.93)	*60.523 (-14.38)*	28.459 (-18.87)	-13.333 (-4.08)	-23.867 (+1.78)
	LN	84.800 (+7.23)	*76.127 (+7.70)*	37.723 (+7.56)	-13.182 (-2.90)	-24.585 (-1.18)
	MC	79.082	*70.686*	35.073	-12.810	-24.299

Finally, one of the important features of Conditional Tail Expectation shall be recalled. Since CTE is a concave risk measure (see Section 5.4 for details), then

$$V^l \leq_{cx} V \leq_{cx} V^c$$

implies corresponding ordering of Conditional Left Tail Expectations of V^l, V and V^c, i.e.

$$Risk_2[V^l] \leq Risk_2[V] \leq Risk_2[V^c] \tag{61}$$

(see auxiliary Formula (49) for details). This fact is justified by numerous numerical results (see Tables 2, 4, 6, 8 and 10). Recall that $Risk_1$ is defined as

$$Risk_1[X] = b - Q_p[X]$$

(see auxiliary Formula (50)). If the convex order would lead to corresponding ordering of quantiles, then the property (61) would be expected to take the form

$$Risk_1[V^c] \leq Risk_1[V] \leq Risk_1[V^l]$$

for the quantile risk measure. It should be noted that this property does not necessarily hold true, since quantile is not a concave risk measure (one can observe that $p = 0.99$ is the only case where it is fulfilled).

9 Short summary and outlook

In this thesis the performances of different approximations for computing quantiles and Conditional Tail Expectations of the stochastic final value of a series of discrete payments were compared for a wide range of parameters. The efficiency of two well-known moment-matching approximations and innovative comonotonic approximations was examined by comparing with the results obtained by Monte Carlo simulation. Numerous investigations which were carried out at varying parameters show that in the case of stochastic final value, generally, lower bound approach leads to more accurate approximations in comparison with "maximal variance" and two moment-matching approximations. It should be remarked, that analogous investigations were carried out by Vanduffel et al. (2005a) in the case of stochastic present value of a series of future deterministic payments. They show that "maximal variance" lower bound approach outperforms both moment-matching approximations.

The "comonotonic" methodology finds much wider range of applications than the ones presented in this thesis. For example, the idea of comonotonic approximation can be extended to the case of so-called "saving-consumption" plan, when positive payments (savings) are followed by negative ones (consumptions). In this case in the contrast to the situation with positive payments, the lower bound will in general not be a comonotonic sum, thus the "distorted expectations" can not be easily obtained by summing the corresponding risk measures of the terms in the sum. Additionally, both moment-matching approximations can be used only in the case of positive cash flows. To overcome this problem Goovaerts et al. (2004) use an approach where convex bounds for positive and negative sums are treated separately. Another idea how to find accurate approximations in this special case is considered in Vanduffel et al. (2005b). Vanduffel et al. (2003) consider the problem how to determine the required level of the current provision to be able to pay a series of deterministic obligations. They also use comonotonic approximations to solve this problem. Ahcan et al. (2006) extended the comonotonic approximations to the case of present values of series of random payments. Simon et al. (2000) and Reynaerts et al. (2006) use comonotonicity technique to price Asian options. Valdez et al. (2003) constructed upper and lower convex bounds for the distribution of a sum of non-independent log-elliptical random variables, which are extensions of the log-normal random variables.

10 References

1. Ahcan A., Darkiewicz G., Goovaerts M., Hoedemakers T. (2006). Computation of convex bounds for present value functions with random payments. Journal of Computational and Applied Mathematics 186, 23-42.

2. Artzner P., Delbaen F., Eber J.-M., Heath D. (1999). Coherent measures of risk. Mathematical Finance 9(3), 203-228.

3. Artzner P. (1999). Application of coherent risk measures to capital requirements in insurance. North American Actuarial Journal 3(2), 11-25.

4. Bäuerle N., Müller A. (1998). Modeling and comparing dependencies in multivariate risk portfolios. ASTIN Bulletin 28(1), 59-76.

5. Deelstra G., Liinev J., Vanmaele M. (2004). Pricing of arithmetic basket options by conditioning. Insurance: Mathematics and Economics 34, 55-77.

6. Denneberg D. (1994). Non-additive measure and integral. Kluwer Academic Publishers, Dordrecht.

7. Dhaene J., Goovaerts M.J. (1995). On the dependency of risks in the individual life model. DTEW Research Report 9539, K.U.Leuven.

8. Dhaene J., Goovaerts M.J. (1996). Dependency of risks and stop-loss order. ASTIN Bulletin 26(2), 201-212.

9. Dhaene J., Wang S., Young V., Goovaerts M.J. (2000). Comonotonicity and maximal stop-loss premiums. Mitteilungen der Schweiz. Actuarvereinigung 2, 99-113.

10. Dhaene J., Denuit M., Goovaerts M.J., Kaas R., Vyncke D. (2002a). The concept of comonotonicity in actuarial science and finance: theory. Insurance: Mathematics and Economics 31, 3-33.

11. Dhaene J., Denuit M., Goovaerts M.J., Kaas R., Vyncke D. (2002b). The concept of comonotonicity in actuarial science and finance: applications. Insurance: Mathematics and Economics 31, 133-161.

12. Dhaene J., Vanduffel S., Tang Q., Goovaerts M.J., Kaas R., Vyncke D. (2004). Solvency capital, risk measures and comonotonicity: a review. DTEW Research Report 0416, K.U.Leuven.

13. Dufresne D. (2004a). Bessel processes and a functional of Brownian motion. Preprint. Center for Actuarial Studies of University of Melbourne.

14. Dufresne D. (2004b). The log-normal approximation in financial and other computations. Advances in Applied Probability 36, 747-773.

15. Goovaerts M.J., Dhaene J. (1999). Supermodular ordering and stochastic annuities. Insurance: Mathematics and Economics 24, 281-290.

16. Goovaerts M.J., Redant H. (1999). On the distribution of IBNR reserves. Insurance: Mathematics and Economics 25, 1-9.

17. Goovaerts M.J., Dhaene J., Schepper A. (2000). Stochastic upper bounds for present value functions. Journal of Risk and Insurance 67(1), 1-14.

18. Goovaerts M.J., De Schepper A., Hua Y., Darkiewicz G., Vyncke D. (2004). An investigation on the use of copulas when calculating general cash flow distributions. Proceedings of the 8th International Congress on Insurance: Mathematics Economics (IME2004), Rome (Italy), Jun. 14-16.

19. Kaas R., Dhaene J., Goovaerts M.J. (2000). Upper and lower bounds for sums of random variables. Insurance: Mathematics and Economics 27, 151-168.

20. Kaas R., Goovaerts M., Dhaene J., Deniut M. (2001). Modern actuarial risk theory. Kluwer Academic Publisher, Dordrecht.

21. Milevsky M.A. (1997). The present value of stochastic perpetuity and the Gamma distribution. Insurance: Mathematics and Economics 20, 243-250.

22. Milevsky M.A., Posner S.E. (1998). Asian options, the sum of lognormals, and reciprocal Gamma distribution. Journal of Financial and Quantitative Analysis 33(3), 409-422.

23. Panjer H.H. (2002). Measurement of risk, solvency requirements and allocation of capital within financial conglomerates. Proceedings of 2002 AFIR (Actuarial Approach to Financial Risks), March 2002, Cancun, Mexico.

24. Reynaerts H., Vanmale M., Dhaene J., Deelstra G. (2006). Bounds for the price of a European-style Asian option in a binary tree model. European Journal of Operational Research 168, 322-332.

25. Röell A. (1987). Risk aversion in Quiggin and Yaari's rank-order model of choice under uncertainty. The Economic Journal 97 (Conference 1987), 143-159.

26. Shiryaev A.N. (1996). Probability. Springer, New York.

27. Simon S., Goovaerts M.J., Dhaene J. (2000). An easy computable upper bound for the price of an arithmetic Asian option. Insurance: Mathematics and Economics 26, 175-183.

28. Valdez E., Dhaene J. (2003). Bounds for sums of dependent log-elliptical risks. Proceedings of the 7th International Congress on Insurance: Mathematics Economics (IME2003), Lyon (France), Jun. 25-27.

29. Vanduffel S., Dhaene J., Goovaerts M.J., Kaas R. (2003). The hurdle-race problem. Insurance: Mathematics Economics 33, 405 - 413.

30. Vanduffel S., Hoedemakers T., Dhaene J. (2005a). Comparing approximations for risk measures of sums of non-independent lognormal random variables. North American Actuarial Journal 9(4), 71-82.

31. Vanduffel S., Dhaene J., Goovaerts M.J. (2005b). On the evaluation of 'saving-consumption' plans. Journal of Pension Economics and Finance 4(1), 17 - 30.

32. Vanmale M., Deelstra G., Liinev J., Dhaene J., Goovaerts M.J. (2006). Bounds for the price of discrete arithmetic Asian option. Journal of Computational and Applied Mathematics 185, 51-90.

33. Vyncke D., Goovaerts M., Dhaene J. (2001). Convex upper and lower bounds for present value functions. Applied Stochastic Models in Business and Industry 17(2), 149-164.

34. Wang S.S. (1996). Premium calculation by transforming the layer premium density. ASTIN Bulletin 26(1), 71-92.

35. Wang S.S., Young V.R. (1998). Ordering risks: Expected utility theory versus Yaari's dual theory of risk. Insurance: Mathematics and Economics 22, 145-161.

36. Wang S.S. (2000). A class of distortion operators for pricing financial and insurance risks. Journal of Risk and Insurance 67(1), 15-36.

37. Wang S.S. (2002). A risk measure that goes beyond coherence. Proceedings of 2002 AFIR (Actuarial Approach to Financial Risks), March 2002, Cancun, Mexico.

38. Yaari M.E. (1987). The dual theory of choice under risk. Econometrica 55(1), 95-115.

39. Yor M. (2001). Exponential functionals of Brownian motion and related processes. Springer, Berlin.

11 Erklärung

Ich erkläre an Eides Statt, daß ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Chemnitz, den 30. November 2006

12 Thesen

1. In this thesis the performances of different approximations are compared for a standard actuarial and financial problem: the estimation of quantiles and conditional tail expectations of the final value of a series of discrete cash flows.

2. To calculate the risk measures such as quantiles and Conditional Tail Expectations, one needs the distribution function of the final wealth. The final value of a series of discrete payments in the considered model is the sum of dependent lognormal random variables. Unfortunately, its distribution function cannot be determined analytically. Thus usually one has to use time-consuming Monte Carlo simulations. Computational time still remains a serious drawback of Monte Carlo simulations, thus several analytical techniques for approximating the distribution function of final wealth are proposed in the frame of this thesis. These are the widely used moment-matching approximations and innovative comonotonic approximations.

3. Moment-matching methods approximate the unknown distribution function by a given one in such a way that some characteristics (in the present case the first two moments) coincide. The ideas of two well-known approximations are described briefly. Analytical formulas for valuing quantiles and Conditional Tail Expectations are derived for both approximations.

4. Recently, a large group of scientists from Catholic University Leuven in Belgium has derived comonotonic upper and comonotonic lower bounds for sums of dependent lognormal random variables. These bounds are bounds in the terms of "convex order". In order to provide the theoretical background for comonotonic approximations several fundamental ordering concepts such as stochastic dominance, stop-loss and convex order and some important relations between them are introduced. The last two concepts are closely related. Both stochastic orders express which of two random variables is the "less dangerous/more attractive" one.

5. The central idea of comonotonic upper bound approximation is to replace the original sum, presenting final wealth, by a new sum, for which the components have the same marginal distributions as the components in the original sum, but with "more dangerous/less attractive" dependence structure. The upper bound, or saying mathematically, convex largest sum is obtained when the components of the sum are the components of comonotonic random vector. Therefore, fundamental concepts of comonotonicity theory which are important for the derivation of convex bounds are introduced. The most wide-spread examples of comonotonicity which emerge in financial context are described.

6. In addition to the upper bound a lower bound can be derived as well. This provides one with a measure of the reliability of the upper bound. The lower bound approach is based on the technique of conditioning. It is obtained by applying Jensen's inequality for conditional expectations to the original sum of dependent random variables. Two slightly different version of conditioning random variable are considered in the context of this thesis. They give rise to two different approaches which are referred to as comonotonic lower bound and comonotonic "maximal variance" lower bound approaches.

7. Special attention is given to the class of distortion risk measures. It is shown that the quantile risk measure as well as Conditional Tail Expectation (under some additional conditions) belong to this class. It is proved that both risk measures being under consideration are additive for a sum of comonotonic random variables, i.e. quantile and Conditional Tail Expectation for a comonotonic upper and lower bounds can easily be obtained by summing the corresponding risk measures of the marginals involved.

8. A special subclass of distortion risk measures which is referred to as class of concave distortion risk measures is also under consideration. It is shown that quantile risk measure is not a concave distortion risk measure while Conditional Tail Expectation (under some additional conditions) is a concave distortion risk measure. A theoretical justification for the fact that "concave" Conditional Tail Expectation preserves convex order relation between random variables is given. It is shown that this property does not necessarily hold for the quantile risk measure, as it is not a concave risk measure.

9. Finally, the accuracy and efficiency of two moment-matching, comonotonic upper bound, comonotonic lower bound and "maximal variance" lower bound approximations are examined for a wide range of parameters by comparing with the results obtained by Monte Carlo simulation. It is justified by numerical results that, generally, in the current situation lower bound approach outperforms other methods. Moreover, the preservation of convex order relation between the convex bounds for the final wealth by Conditional Tail Expectation is demonstrated by numerical results. It is justified numerically that this property does not necessarily hold true for the quantile.

13 Acknowledgements

With a deep gratitude I wish to express my sincere thank to my supervisor Dr. Matthias Richter. He carefully accompanied all my work and encouraged me. His explanations were very lively and easy understandable. I will remember all my life his irreproachable supervision and help.

I am thankful to Prof. Dr. Ralf Wunderlich for discussions and useful advices.

My sincere and enormous thanks are to my best friends Anna Mogilatenko and Igor Afonin for the support and help.

I feel a deep gratitude to my mother Natalia Gendruseva who formed part of my vision and taught me the good and useful things.

I am also grateful to my sister Zinaida Gendruseva for sisterhood and moral support.

I would like to thank my father-in-law Boris Karniychuk and mother-in-law Galina Karniychuk for their help, care and comprehension, which our family were surrounded.

Especially, I would like to thank my husband Maksim and daughter Polina for the inspiration and moral support. Without their loving and understanding I would never complete my work.

For the financial support I would like to thank Karl & Ruth Fund and German Academic Exchange Service (DAAD).